

Université du Québec
Institut national de la recherche scientifique
Centre Énergie Matériaux et Télécommunications

**Amélioration de l'Efficacité Énergétique des Futurs Réseaux Sans Fil
Hétérogènes par Partition de Graphes**

Par
Salma Fazazi

Mémoire présenté pour l'obtention du grade de
Maître es Sciences, M.Sc.
en Mémoire présenté pour l'obtention du grade de

Jury d'évaluation

Examineur externe	Nadir Hakem UQAT
Examineur interne	Tarek Djerafi Centre EMT, INRS
Directeur de recherche	Sofiène Affes Centre EMT, INRS

REMERCIEMENTS

Mes premiers mots de remerciements ne sauraient être destinés à autre que M. Sofène Affes. Ce manager hors pair à l'expérience accrue est doté d'une bienveillance rare et d'un dynamisme entraînant. Je le remercie d'abord de m'avoir recrutée et de m'avoir fait confiance dès le premier jour me poussant à me surpasser et à monter en compétence. Je le remercie ensuite pour sa gentillesse inouïe et lui souhaite une très bonne continuation dans ses fonctions. Je tiens à remercier également M. Oussama Ben Smida pour sa patience et son aide précieuse pendant la durée de ma maîtrise. Sa présence agréable est aussi une de ses qualités pour lesquelles il est apprécié de tous. Et que serait le plateau du sixième étage sans la bonne humeur contagieuse de l'équipe Wireless Lab. Je les remercie chaleureusement pour leur accueil bienveillant, ma maîtrise n'aurait pas été la même sans leur présence.

RÉSUMÉ

Comme le secteur des télécommunications mobiles se dirige vers la 5G, la quantité de données traversant les réseaux sera en augmentation perpétuelle. De ce fait, les opérateurs quels qu'ils soient, se retrouvent dans la nécessité de répondre à cette croissance en transformant le réseau d'accès radio (RAN), afin de garantir la disponibilité et la qualité de leurs services. Sauf que nous nous retrouvons dans une éventualité où la consommation d'énergie risque d'accroître vigoureusement.

C'est bien dans ce sens que notre travail visera à proposer une méthode où les stations de base seront dans l'une des deux états (activé,désactivé), tout en garantissant la qualité de services des usagers, dans le but de réduire la consommation d'énergie totale du réseau.

L'intérêt suscité par ce concept ne date pas d'hier bien que la littérature sur le sujet est pour le moins abondante. Sauf que ces travaux réalisés abordent le problème, qui est NP-difficile, d'un point de vue d'optimization mathématique. De plus, si peu ont été établis pour le modèle Hétérogène. En effet, les auteurs de [5] considèrent une charge dynamique de trafic cellulaire et formulent un algorithme de *bin packing* combiné avec un algorithme heuristique, *simulated annealing* (HSA), pour établir un schéma de virtualisation des BBU qui minimise la consommation d'énergie. Cependant, ils n'ont pas pris compte des stations de base traditionnelles qui augmentent le temps de calcul exponentiellement. Par conséquent, nous souhaitons simplifier encore plus notre solution et résoudre le problème en utilisant des algorithmes de partitionnement de graphes où les sorties seraient un ensemble de station de base actif/éteint.

Mots-clés Efficacité énergétique, partitionnement de graphe, recherche tabou, réseau hétérogène.

ABSTRACT

As the mobile telecommunications industry barrels toward the HetNets, the range of the amount of data traffic traversing networks is escalating; which increases both their power consumption and their carbon footprint. Thus, creating a challenge for researchers and industrials to find/propose eco-friendly solutions and green alternatives. In retrospect, multiple papers strived to propose energy saving solutions. As a matter of fact, what all these papers have in common is that they tackled this problem, which is NP-hard, from a mathematical optimization standpoint.

Nevertheless, there had been little work establishing a framework for heterogeneous C-RAN power model. The authors in [3], consider a dynamic cell traffic load and formulates a bin packing algorithm combined with a heuristic simulated annealing algorithm (HSA) to establish a BBUs virtualization scheme that minimizes the power consumption. However, they did not take into consideration the traditional base stations which, in hindsight, made the computational time exponential to the number of the BSs. Howbeit, we would like to simplify even further our solution and tackle the issue using graph partitioning algorithms where the outputs would be a set of active and switched off base stations.

Keywords Energy efficiency, heterogenous network, graph partitioning algorithm, Tabu search.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES FIGURES	xi
LISTE DES TABLEAUX	xiii
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES	1
1 INTRODUCTION GÉNÉRALE	3
1.1 Motivation et contexte	3
1.2 Contributions du mémoire	4
2 REVUE DE L'ÉTAT D'ART	7
2.1 Introduction	7
2.2 L'évolution des réseaux mobiles	7
2.2.1 Architecture du réseau HetNet	8
2.3 Modèle de consommation de puissance	10
2.4 Modèle de puissance en mode désactivé	14
2.5 Conclusion	16
3 PARTITIONNEMENT DE GRAPHERS	17
3.1 Introduction	17
3.2 La problématique partitionnement de graphes	17
3.3 Les approches Heuristiques	18
3.3.1 Kernighan-Lin	18
3.3.2 Fedducia-Mathesis	20
3.3.3 Spectral Clustering	21
3.3.4 La recherche Taboue	25
3.4 Conclusion	27
4 APPROCHE PROPOSÉE DE GESTION DE CONSOMMATION D'ÉNERGIE	29
4.1 Introduction	29
4.2 Description du Simulateur	29

4.2.1	Structure du Simulateur	31
4.2.2	Paramètres de simulation	38
4.3	Conclusion	39
5	RÉSULTATS ET PERFORMANCES	41
5.1	Conclusion	52
6	CONCLUSION	55
	BIBLIOGRAPHIE	57

LISTE DES FIGURES

2.1	Exemple d'un déploiement HetNet	10
2.2	L'effet de la profondeur de désactivation sur une heure de P_{sleep}	16
4.1	Structure du simulateur	31
4.2	Les puissances des signaux d'interférences et le signal reçu par les utilisateurs.	36
4.3	L'algorithme Tabou	37
5.1	Exemple du déploiement du réseau	41
5.2	Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule, et 30 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).	43
5.3	Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule et 50 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).	45
5.4	Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule, et 100 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).	46
5.5	Gains optimaux en puissance et en QoS établis pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L	47
5.6	Fonction utilité optimale 4.1 établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L	48
5.7	Gains optimaux en puissance et QoS établis pour différentes densités en nombre d'utilisateurs par cellule en fonction de la mobilité.	49
5.8	Fonction utilité optimale 4.1 établie pour plusieurs valeurs de w_p en fonction de la mobilité.	50
5.9	Complexité en fonction de la densité en termes du nombre d'utilisateurs par cellule.	52

LISTE DES TABLEAUX

- 2.1 Tableau des caractéristiques de quelques stations des bases 9
- 4.1 Tableau des paramètres de simulation 38
- 4.2 Tableau des paramètres de la recherche taboue 39

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

BS	Station de base
BBU	Unité de bande de base
CAPEX	Capital Expenditures
EE	Efficacité Energétique
GSM	Système mondial pour les communications mobiles
HCRAN	Heterogenous Cloud Radio Access Network
HetNet	Réseau hétérogène
IoT	Internet des objets KPI
Indicateur de performance clé	
LTE	Évolution à long terme
MBS	macro cellule
MIMO	Multiple input multiple output
OPEX	Operational Expenditures
PA	puissance d'amplification
QoS	Qualité du service
RAN	Radio Access Network
RRH	Remote Radio Head
RF	Radio fréquence
RSRP	Puissance reçue
SBS	Petite cellule

TS	Time slot
TTI	Transmission Time Interval
TX	Transmission
UE	Equipement utilisateur
VoLTE	Voice over LTE
1G	Première génération
2G	Deuxième génération
3G	Troisième génération
4G	Quatrième génération
5G	Cinquième génération

Chapitre 1

INTRODUCTION GÉNÉRALE

1.1 Motivation et contexte

La qualité de service a toujours été un souci primordial chez tout opérateur. Depuis, les nouvelles technologies convergent vers l'internet des objets (IoT) où tout est connecté, ce qui a résulté en une augmentation surprenante du trafic. C'est dans ce contexte que le réseau hétérogène C-RAN (ou Cloud RAN), également appelé RAN centralisé, est né. En effet, il est considéré comme l'une des technologies pré-requises permettant le déploiement du réseau à haute performance 5G.

L'architecture HC-RAN est une version améliorée du réseau traditionnel, où on intègre le concept de cloud computing aux systèmes mobiles en toute coexistence avec ce dernier. HC-RAN se compose de nombreuses entités appelées Remote Radio Head (RRH) et d'un pool d'unités de basebande (BBU). Les RRH représentent toute cellule de petite à moyenne couverture tels que Picocells, femtocells ou points d'accès à faible puissance. Ils sont distribués et connectés au pool BBU via des fibres optiques à large bande passante. Ce dernier représente le centre de données hébergeant les serveurs BBU.

L'une des caractéristiques du réseau hétérogène est la densification du réseau. On déploie intensivement une variété de stations de base (femtocell, picocell, MBS, SBS, etc.) afin d'assurer une meilleure qualité de service. Parce que le déploiement dense des stations de base et des points d'accès à petite échelle signifie une réduction du délai de service et un meilleur rapport signal sur bruit et interférences (SNIR). Ceci garantit aux opérateurs une réduction des dépenses en capital

(CAPEX) et des coûts d'exploitation (OPEX) puisqu'on réutilise les ressources et les infrastructures existantes. Ils simplifient donc les opérations et la gestion du réseau et augmentent la capacité et la flexibilité du réseau tout en réduisant la maintenance et la budgétisation de location. De plus, il permet aux opérateurs de co-exécuter les fonctionnalités RAN ainsi que d'autres fonctions réseau dans le centre de données qui réduit le coût total nécessaire à l'exploitation d'un réseau. En fait, pour intégrer les services coopératifs dans les zones denses, il a été proposé que le C-RAN intègre les réseaux macro-cellulaires existants pour créer un réseau d'accès hétérogène.

Toutefois, l'agrégation de la puissance consommée par tous ces composants devient beaucoup plus importante. Par conséquent, plusieurs recherches ont été établies pour améliorer l'efficacité énergétique du réseau. En effet, l'algorithme de gestion de la commutation de station de base (BS) entre l'état actif/désactif est parmi les solutions assurant l'économie d'énergie les plus efficaces.

1.2 Contributions du mémoire

Le but de cette recherche est d'améliorer l'efficacité énergétique du réseau hétérogène en proposant un schéma d'activation / désactivation de stations de base tout en maintenant la qualité de service (QoS). Par conséquent, la stratégie d'activation/désactivation des BS doit être appliquée de manière dynamique tout en tenant compte du compromis entre l'économie d'énergie et la qualité de service des utilisateurs attachées aux cellules désactivées. En effet, dans [10], ils proposent de regrouper les stations de base où chaque groupe de BS vise à minimiser une fonction de coût. Cette fonction est proportionnelle à la consommation d'énergie individuelle des stations de bases et le temps nécessaire du transfert de leurs charge. Ces clusters sont non coopératives et , par conséquent, ils formulent le problème comme un jeu non coopératif parmi les clusters. Pour résoudre ce jeu, ils proposent un algorithme distribué en utilisant des notions d'apprentissage du regret, dans lesquelles les clusters implicitement coordonnent leurs transmissions sans échanger leurs informations. Dans [9], la stratégie d'activation/de désactivation intra-cluster est choisie en fonction de l'impacte de la charge sur les autres BS du cluster. Pareil pour [11] et [30], où ils ont entrepris d'optimiser l'allocation des ressources afin d'impacter la distribution de la charge et éventuellement réduire la consommation d'énergie dans un BS. En effet, [30] ont proposé de décomposer la problématique en deux sous-problème (association des UEs et commutation des BSs), puis ils ont choisi une méthode greedy qui définit une fonction de coût qui est proportionnelle à un KPI de la QoS (délai). De plus,

le modèle de puissance de [31] est fixé de manière non proportionnelle à la charge de trafic, mais leur solution associe la planification de commutation des BS et l'association des utilisateurs, afin de maximiser la métrique EE d'un MIMO HetNet massif tout en respectant les contraintes de charge de trafic de tous les BS. Contrairement aux modèles précédents, [31] approche le problème avec une métrique équilibrant le compromis entre le débit et la consommation d'énergie. Tandis que [32] a mis l'accent sur la fiabilité du service, proposant ainsi le concept du sommeil opportuniste, dans lequel un BS est en mode sommeil que pendant certaines périodes. En outre, La planification BS on/off et l'allocation des ressources sont conjointement considérées pour minimiser la consommation d'énergie du réseau, sous contraintes de la probabilité moyenne de blocage des utilisateurs. Ainsi, un jeu d'enchères entre les utilisateurs et les BS est formulé, et le résultat initial du jeu détermine une stratégie d'association d'utilisateurs. Quand les associations d'utilisateurs sont déterminées, chaque BS prend sa propre décision concernant la commutation on/off selon la fiabilité du service.

Dès lors, une solution, à but d'économiser l'énergie globale, pourrait être pénalisante pour les utilisateurs locaux en terme de qualité des services. Par conséquent, nous décrivons systématiquement le modèle de puissance des systèmes cellulaires HetNet pour en déduire un algorithme de sauvegarde qui tient en compte l'association des utilisateurs, de la mise en cluster et d'un schéma d'activation / désactivation. Dans [10] et [6], ils ont utilisé le clustering pour créer des stratégies de transmission optimales minimisant les dépenses énergétiques. Cependant, cela conduit à un compromis entre la réduction d'énergie et les performances de débit induites par la compétition intercluster.

Pour répondre à notre problématique, nous adoptant le plan suivant: nous présentons l'architecture du réseau HetNet dans le chapitre II. Puis nous abordons la problématique traitée ainsi que le modèle de puissance qui servent de base à notre étude dans le chapitre III. Dans la chapitre IV, nous discutons la théorie de partitionnement de graphe fondamentale à notre approche. Ensuite, nous examinons la solution proposée et le simulateur hebergeant notre étude, et enfin nous analysons les résultats retenus dans chapitre V.

Chapitre 2

REVUE DE L'ÉTAT D'ART

2.1 Introduction

Au cours des deux dernières années, le milieu de la recherche s'est tourné vers les solutions d'économie d'énergie et d'efficacité énergétique. Les chercheurs ont commencé à cibler des projets tels que GreenTouch, Energy Aware Radio and network technology (EARTH) et Towards Real and Energy Efficient Network Design (TREND) pour un réseau cellulaire, 5G, vert. De plus, l'introduction du réseau hétérogène Cloud RAN permet d'atteindre des gains en terme de capacité ainsi que l'efficacité énergétique. En fait, l'analyse de la consommation d'énergie du réseau cellulaire mobile indique qu'environ 80% de l'énergie totale est consommée par les stations de base (BS) [1]. En effet, une station de base consomme une certaine quantité d'énergie pour maintenir son fonctionnement normal, notamment de l'énergie pour ses circuits, son système de refroidissement, etc...

Evidemment, dans ce chapitre nous présentons une vue d'ensemble l'architecture du réseau hétérogène avant d'évaluer le modèle de puissance de chaque composante.

2.2 L'évolution des réseaux mobiles

À présent, les réseaux mobiles sont à leur cinquième génération. La première génération comprend des systèmes de communications analogiques dédiés à la transmission de la voix. Ce système reconnaît plusieurs points faibles, à savoir la limitation de la capacité et de la mobilité, et l'absence

du mécanisme de sécurité. La deuxième génération était conçue pour surmonter les faiblesses de la 1G, ainsi qu'être en phase avec les applications multimédias et les services à haute vitesse. Tandis que la troisième génération était déployée pour assurer l'interopérabilité de tous les systèmes de communications câblées et sans fil afin de garantir une mobilité globale et un accès aux services sans coupure. Conformément, les systèmes 3G sont caractérisés par leur souplesse et leur meilleure capacité.

La quatrième génération(4G) est établie pour accomoder les transfert de données à très haut débit et attribuer plus d'autonomie aux stations de base. La 4G est un réseau tout IP: la transmission et la signalisation se font à travers des paquets IP, ainsi le réseau coeur d'un système LTE n'inclut que le domaine paquet et on peut parler d'une connectivité permanente tout-IP. De plus, la notion de la qualité de service (QoS) est intégrée en se basant sur des critères tels que : la priorité, la fiabilité, le délai, et le débit. Cependant, la technologie LTE était conçue pour la transmission des données à très haut débit.

Les réseaux de la cinquième générations(5G) sont conçus pour avoir accès à des très haut debits (de plusieurs gigabits par seconde) avec haute disponibilité et faible latence, ainsi que l'implémentation massive des communications de type machines. Cette technologie inspire à réaliser une connectivité totale qui est l'essor du concept de ville intelligente.

2.2.1 Architecture du réseau HetNet

Nous avons mentionné précédemment que le réseau 5G vise à établir une connectivité totale, qui induit une croissance exponentielle du trafic de données. Et afin que le réseau puisse satisfaire cette demande en termes de capacité et de débit, il fallait que le réseau modifie son infrastructure. De ce fait, un déploiement dense de plusieurs stations de bases de différents rayons de couverture, s'est avéré rentable. En effet, vue les coûts énormes d'installation d'une station de base, plusieurs opérateurs choisissent de réutiliser leurs ressources en termes des station de bases traditionnelles. Ces stations de bases sont de grandes couvertures, appelées macrocellules. Cette dernière est caractérisée par une puissance de rayonnement énorme et unifrome. Tandis que les stations de bases picocellules et femtocellules ont des distances de couverture moyenne à petite.

- Macrocell : sont des cellules à puissance de rayonnement large , offrant une couverture qui étend à quelques kilomètres et desservent un grand nombre d'utilisateurs . En général, ces BS sont installés dans les zones urbaines et rurales .
- Picocell :sont des cellules à puissance de rayonnement moyenne , offrant une couverture qui s'étend à des centaines de mètres et desservent des dizaines d'utilisateurs . En général, ces BS sont installés dans les environnements indoor comme les immeubles et les bureaux.
- Femtocell :sont des cellules à puissance de rayonnement petit , offrant une couverture qui s'étend à des dizaines de mètres et desservent quelques d'utilisateurs . En général, ces BS sont installés dans les environnements indoor comme les maisons et les entreprises, ou dans un environnement outdoor comme un espace publique.

Le tableau suivant,2.1, présente quelques caractéristiques des BS macro et pico.

Paramètres	Macrocell	Picocell	Femtocell
Puissance de transmission	43 dBm	23-30 dBm	20 dBm
Rayon de couverture	Quelques kilomètres	100m-300m	10m-20m
Nombre d'utilisateurs	3000	20-30	4-6

Tableau 2.1 – Tableau des caractéristiques de quelques stations des bases

Le réseux déployant toutes ces différentes BS est nommé réseau hétérogène ou HetNet. Il assure une extension de couverture et une amélioration de débit. La figure 2.1 illustre en gros l'architecture d'un tel réseau.

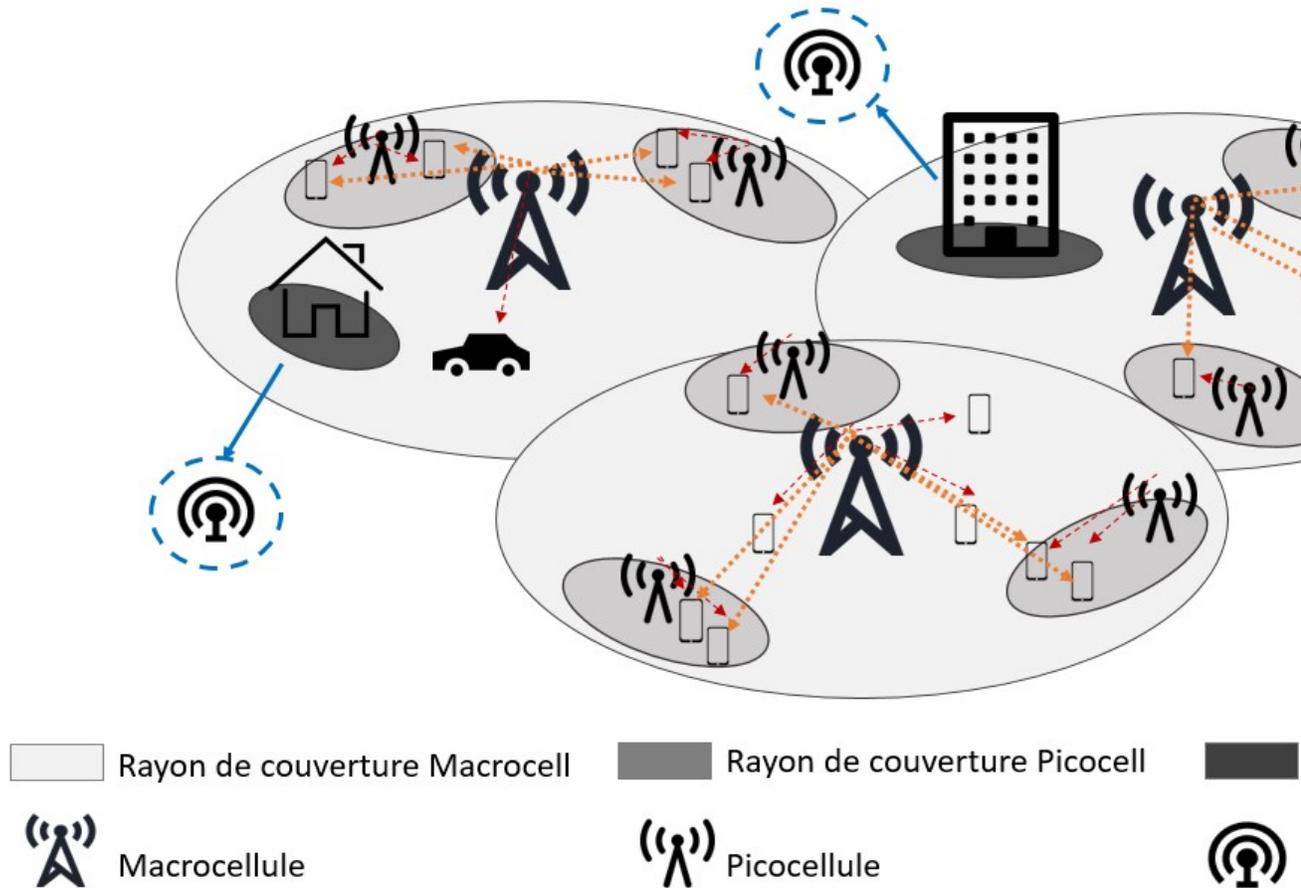


Figure 2.1 – Exemple d'un déploiement HetNet

2.3 Modèle de consommation de puissance

L'objectif de cette recherche est de minimiser la consommation d'énergie du réseau H-CRAN. Dans ce cas, en modélisant la consommation électrique de l'ensemble du réseau, on comprend mieux quel composant consomme plus d'énergie que d'autre, quelle est l'unité la plus affectée par les ressources et quel est l'impact sur le réseau en terme d'efficacité énergétique. Ainsi, la modélisation au niveau du système nous permettra de mieux maîtriser la stratégie d'activation/désactivation.

En considérant l'architecture HetNet (voir figure 2.1), la modélisation de la cellule macro (MBS) et pico sera la même. Dans [17], l'auteur décrit la consommation électrique du réseau à l'aide d'un modèle paramétré qui prend en compte plusieurs facteurs tels que la largeur de bande, le nombre d'antennes, les secteurs macro-BS, et les RRH variables. Sauf que, quelle que soit la granularité de leur modèle, ceci rend le schéma de commutation on/off beaucoup plus complexe. De même, ils n'ont pas considéré que la réalité fait que le réseau soit hétérogène; vu qu'on met au service les stations de bases MBS et SBS. Par conséquent, nous fournissons un modèle qui traite exclusivement cette architecture.

La BS est responsable de la génération de signaux numériques faite au niveau basebande(BB), puis de la transmission aux utilisateurs à travers l'émetteur-récepteur RF. Avec la technologie du facteur d'échelle, nous pouvons calculer le nombre d'opérations de la BS au niveau BB par seconde par watt ou en opération par unité (GOPS)/watt, voir l'équation (2.1). En fait, [2] décrit un modèle de puissance générique qui s'adapte au matériel de la station de base(type de BS, nombre d'antennes, largeur de la bande maximale, etc.), ainsi qu'à sa configuration (charge de trafic, stratégies d'économie d'énergie, nombre d'antennes activées, etc.). Et ils expliquent comment le facteur d'échelle est utilisé dans le modèle. En bref, la puissance consommée par les composants est dérivée d'une valeur de puissance de référence, qui est définie par le fournisseur et combinée à un certain nombre de règles de mise à l'échelle. Ceci nous permettra de mesurer correctement la valeur réelle de la consommation d'énergie lors de la modification d'un paramètre tel que (largeur de bande ou nombre d'antennes). Le document fournit plusieurs tableaux de valeurs de consommation d'énergie de références et de facteurs d'échelle. Selon [2], il existe six paramètres qui influencent la consommation d'énergie de toute station de base : bande passante, efficacité spectrale (qui est essentiellement le taux de codage et l'ordre de constellation), le nombre d'antennes, le domaine de fréquence, et la quantification.

$$P_{BB} = \sum_{i \in I} P_{(i, BB)}^{ref} * \prod_{x \in X} \left(\frac{x_{ac}}{x_{ref}} \right)^{s_{i,x}} \quad (2.1)$$

ou I est l'ensemble des sous-composants de la station de base et X est l'ensemble des paramètres x discutés dans précédemment. $P_{(i, BB)}^{ref}$ est la puissance de consommation des composants en charge de la fonction basebande, tandis que x_{ac} est la valeur actuelle de la puissance consommée par les paramètres x . x_{ref} représente la valeur de puissance de référence, $s_{i,x}$ est le vecteur

d'échelle remis soit à 0 ou 1. Ce dernier indique si la puissance de consommation du sous-composant i est en fonction du paramètre x ou pas.

En outre, la station de base nécessite une puissance continue adéquate pour maintenir son fonctionnement typique. Par conséquent, nous devrions prendre en compte la consommation électrique de la conversion continu-continu qui est formulé comme suit:

$$P_{DC,DC} = \sum_{b \in M} l(\eta_{DC,DC})P_{BB} \quad (2.2)$$

Où $l(\eta_{DC,DC})$ est la fonction de perte du rendement de conversion en courant continu, ce qui prouve que la consommation d'énergie du convertisseur est proportionnelle linéairement à la puissance de consommation des composants de la BS et M denote l'ensemble des BS actives dans le réseau. De plus, cette tension continue provient d'une source d'alimentation AC. Par conséquent, nous devons tenir compte de la consommation d'énergie des convertisseurs AC-DC, ce qui se traduit par:

$$P_{AC,DC} = \sum_{b \in M} l(\eta_{AC,DC}) * (P_{DC,DC} + P_{BB}) \quad (2.3)$$

Lorsque tous ces composants chauffent, il est nécessaire de configurer une unité de refroidissement actif pour le hardware des BS, le convertisseur AC-DC et le convertisseur AC-AC. Ce dernier est généralement l'unité qui consomme le plus d'énergie, et il est proportionnel à la consommation d'énergie de tous les autres composants:

$$P_{cool} = \sum_{b \in M} l_{cool} * (P_{AC,DC} + P_{DC,DC} + P_{BB}) \quad (2.4)$$

Dès lors, les paramètres des stations de base sont déjà définis, la consommation d'énergie du matériel étant constante. Dans la suite de mémoire, nous la prendrons comme telle et lui assignerons P_{static} .

Par ailleurs, la station de base se compose aussi de la partie radio (RF) dont la puissance de consommation correspond aux composants d'amplification PA P_{PA} et radio RF P_{RF} . En effet, le composant qui consomme le plus est l'amplification qui dépend de la puissance de sortie, qu'elle-même dépend des besoins de la distance UE-BS et le débit de données de cet utilisateur, [9]. En

conclusion, nous définissons la puissance de consommation d'une BS comme suit:

$$P_{BS} = P_{RF} * N_{antenna} + P_{PA} + P_{static} \quad (2.5)$$

Cependant, la puissance d'amplification PA (c'est la puissance consommé par un amplificateur) P_{PA} d'une station de base b_1 dépend de la charge de la station de base et le coefficient d'amplification η_{PA} suivant l'équation suivante:

$$P_{PA} = \frac{P_{b_1}^{TX} * \rho_{b_1}}{\eta_{PA}} \quad (2.6)$$

où $P_{b_1}^{TX}$ démontre la puissance de transmission par unité de ressources et ρ_{b_1} le nombre d'unités de ressources nécessaires pour atteindre le débit de données de l'utilisateur. Supposons que chaque UE k est desservi par b_1 . L'utilisateur requiert un débit λ_k . Et supposons que $d_{b_1,k}$ est la distance entre UE et BS, β est le facteur pathloss puis $h_{b_1,k}$ représente l'effet fading. Par conséquent, le débit reçu est calculé comme suit :

$$r_{b_1,k} = B * \text{Log}_2\left(1 + \frac{P_{b_1,k}^{TX} * d_{b_1,k}^\beta * h_{b_1,k}}{N_0 * B + I}\right) \quad (2.7)$$

Sachant que N_0 est le bruit thermique, I sont les interférences et B la largeur de bande. Pour que la station de base garantie une bonne QoS, il faut que la contrainte suivante soit assurée:

$$r_{b_1,k} \geq \lambda_k \quad (2.8)$$

Pour minimiser la consommation d'énergie du système, nous avons besoin d'une allocation optimale des ressources au niveau de l'unité centrale afin d'éviter le gaspillage de ressources. Le document [23] proposait un algorithme permettant de calculer le taux d'utilisation et de déterminer le seuil optimal à partir duquel l'activation/désactivation devrait être appliquée. Rétrospectivement, nous proposons d'appliquer la formation de groupes basée sur la dissimilarité de charge afin d'augmenter notre probabilité de trouver une combinaison optimale. Selon le modèle EARTH, la consommation électrique d'un serveur en mode inactif ne diminue pas car tous les composants du serveur sont toujours allumés. Par conséquent, puisque nous avons l'intention d'éteindre le serveur (i.e., certains composants de la BS), il convient seulement d'inclure la puissance de réveil du serveur. Certains travaux [13] et [2] considèrent que la puissance de réveil est dynamique, proportionnelle à la période de veille du serveur, qui varie de quelques microsecondes à quelques secondes. En fait, plus le temps

de sommeil est long, plus le nombre de composants désactivés est important, ce qui se traduit par un besoin accru de puissance de réveil. Sur ce, la consommation totale d'une BS b_1 pendant une période donnée t est:

$$P_{b_1,t} = a_{b_1,t} * P_{b_1} + |a_{b_1,t} - a_{b_1,t-1}| * P^{sleep} \quad (2.9)$$

où

$$a_{b_1,t} = \begin{cases} 1, & \text{si la station de base } b_1 \text{ était active durant la période } t \\ 0, & \text{si la station de base } b_1 \text{ était inactive durant la période } t \end{cases} \quad (2.10)$$

2.4 Modèle de puissance en mode désactivé

Dans ce contexte, la puissance consommée pendant la période de veille correspond à la puissance consommée par les composants actifs pendant cette période. En fait, pour que le serveur puisse être activé, il ne doit pas être complètement éteint. il peut encore consommer une certaine quantité d'énergie, telle que la puissance de détection [17]. Dans [13], on modélise les activation/désactivation et la consommation d'énergie au réveil de tous les types de stations de base en fonction de la profondeur de sommeil¹ (période de désactivation réelle) dans lesquels ils définissent les modes de sommeil selon quatre états correspondant à la profondeur du désactivation: $71\eta s$, $1ms$, $10ms$ et $1s$. L'objectif de leur travail était de trouver la durée de désactivation optimale correspondant au profil de trafic; ce qui implique que la désactivation et le changement d'états ne doivent pas dépasser la période pendant laquelle la station de base est en mode veille, tout en réalisant une réduction de l'énergie. Cela pourrait être traduit dans les contraintes (2.13) et (2.14).

Supposons que nous avons une connaissance préalable du trafic quotidien moyen. Le processus d'arrivée du trafic est donc un processus de Poisson avec un taux d'arrivée de λ et un temps de service de h_B . De plus, le temps actif de BS correspond à la distribution des heures d'arrivée. Supposons aussi que τ est la profondeur du désactivation en seconde, eq.2.15. Selon le document, la consommation moyenne d'énergie de la station de base correspond à:

$$E[P] = \frac{E[L_{active}]P_{active} + E[L_{deac}]P_{deac} + E[L_{sleep}]P_{sleep} + E[L_{réveil}]P_{réveil}}{L_{active} + L_{deac} + L_{sleep} + L_{réveil}} \quad (2.11)$$

1. Le terme sommeil et désactivation sont interchangeable.

Où L_{actif} , $L_{désac}$, $L_{désactivation}$, $L_{réveil}$ correspond respectivement à la durée du service actif, à la désactivation, au temps de désactivation et au réveil effectif. Et selon le travail [13], la puissance consommé pendant le sommeil est modélisé comme suit:

$$P_{sleep}(\tau) = P_m \exp((-\omega_1 \log_{10}(\tau * 10^6))^{\frac{1}{\omega_2}}) + constant; \quad (2.12)$$

Bien que, de manière réaliste, le temps de désactivation devrait être supérieur à la latence de la transition (2.13) et de manière optimale, la puissance consommée pendant la latence de la transition et la période de veille ne devrait pas être supérieure à la puissance consommée en mode veille (2.16).

$$\theta = \frac{L_{réveil}}{L_{réveil} + L_{deac}} > 1, \quad (2.13)$$

$$\eta = \frac{L_{réveil}}{L_{deac}} > 1, \quad (2.14)$$

$$\tau = L_{réveil} + L_{deac} + L_{sleep}, \quad (2.15)$$

$$P_{deac} + P_{sleep} + P_{wakeup} < P_{idle}. \quad (2.16)$$

Les résultats numériques montrent que s'ils augmentaient, ils réduisent la consommation moyenne d'énergie, puisqu'ils passent plus de temps à dormir, ce qui aurait donc un impact marginal. Cependant, si nous augmentons le temps de réveil par rapport à la désactivation ou au sommeil, la consommation de puissance moyenne augmente. Cela suggère que nous devrions être prudents lors de la sélection des composants à activer/désactiver lorsque le temps de sommeil est dans la plage du temps de configuration (τ and η)

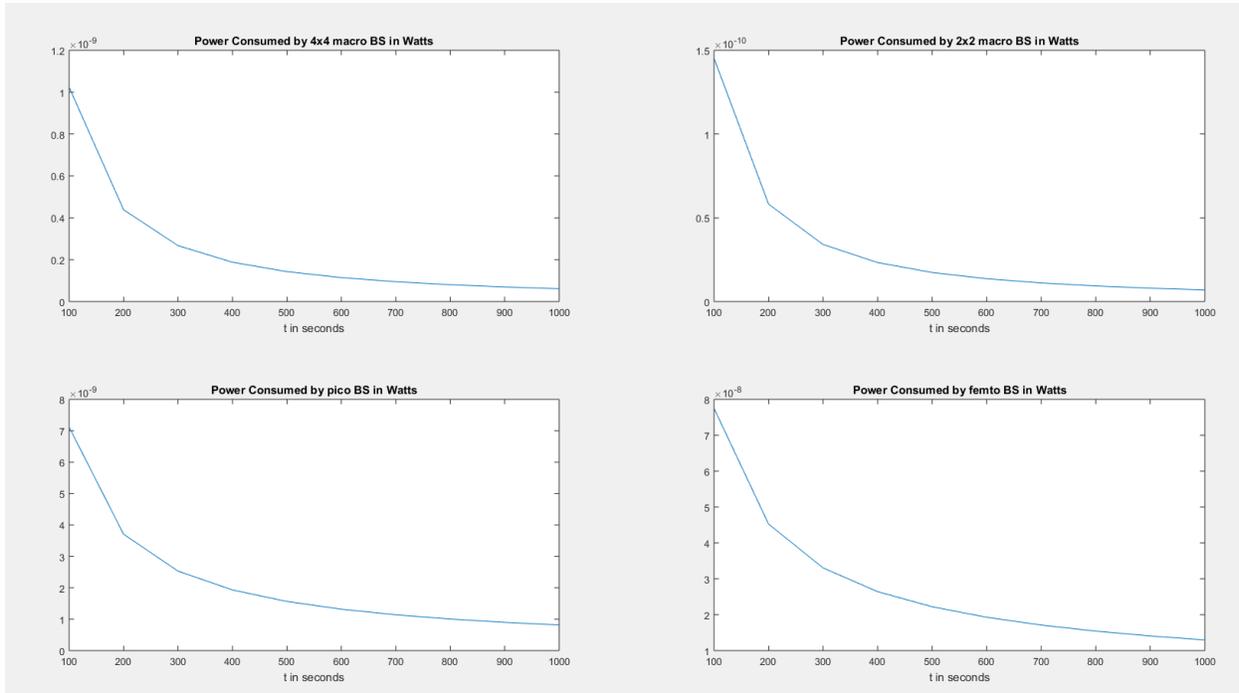


Figure 2.2 – L’effet de la profondeur de désactivation sur une heure de P_{sleep}

Cependant, en réalité et selon un profil de trafic de réseau cellulaire fourni par China Mobile [5], les données recueillies sur le trafic sont qualifiées en moyenne par heure. Ainsi, dans notre cas, une période d’inactivité ou une période de faible trafic s’étend sur des heures. Par conséquent, conformément à l’équation (2.11) et à la figure (2.2), les plages de variation de la consommation d’énergie pendant le mode veille peuvent être considérées comme nulles.

2.5 Conclusion

Dans ce chapitre, nous avons défini et présenté les prérequis nécessaires à la suite de notre étude. Dans le chapitre suivant, nous présentons l’algorithme de partitionnement de graphe qui a inspiré notre stratégie de commutation activation/désactivation des station de bases (BS).

Chapitre 3

PARTITIONNEMENT DE GRAPHES

3.1 Introduction

Dans la théorie des graphes, le partitionnement est l'une des tâches importantes et problématiques. Il consiste à trouver une partition des sommets respectant une ou plusieurs propriétés. Le but de ce chapitre est de présenter une revue de littérature des algorithmes de partitionnement les plus répandus.

3.2 La problématique partitionnement de graphes

Soit le graphe $G = (V, E)$, où V est un ensemble de nœuds et E est un ensemble des arêtes pondérées. L'objectif est de retrouver k sous-ensembles V_1, \dots, V_k tels que:

- $\cup_{i=1}^k V_k = V$ et $V_i \cap V_j = \emptyset$ pour $i \neq j$;
- La somme des arêtes croisant les sous-ensembles est minimisée, i.e, $\min(\text{cutsize}) = \min(\sum_{v \in V_1} \sum_{u \in V_2} E(v, u))$.

3.3 Les approches Heuristiques

3.3.1 Kernighan-Lin

Kernighan – Lin est parmi les algorithmes les plus connus et les plus attrayants pour le partitionnement de graphes vue sa simplicité de déploiement [20]. La première étape consiste à diviser le réseau de manière aléatoire en deux groupes V_1 et V_2 de tailles équitables. La deuxième étape vise à trouver la paire $(i \in V_1, j \in V_2)$ dont l'échange réduirait le plus la taille de coupe, c'est-à-dire le nombre d'arêtes dont les extrémités ne sont pas dans le même groupe, entre V_1 et V_2 . Dans le cas où aucune paire ne réduit la taille de coupe, on cherche la paire qui engendre la plus petite augmentation de cette dernière. A la fin de cette étape, la paire est bloquée. Le processus est ensuite répété jusqu'à ce qu'aucune amélioration supplémentaire ne puisse être obtenue. Tandis que l'algorithme s'exécute jusqu'à ce que toutes les paires soient échangées, nous passons par les états précédents et choisissons parmi eux l'état dans lequel la taille de coupe était la plus petite.

Il est important de souligner le fait que le nombre et la taille des groupes sont fixes, ce qui limite notre stratégie d'activation/désactivation car nous n'avons aucune spécification sur le nombre total de stations de base actives ou inactives. De plus, le processus a été conçu pour préserver une bisection équilibrée, ce qui est un inconvénient dans notre cas.

De plus, le caractère aléatoire de l'attribution initiale peut, bien que pas nécessairement, produire des résultats différents. Ceci dit, le principal inconvénient de cet algorithme est son temps d'exécution, qui est $O(n^3)$, où n est le nombre de sommets.

Si nous fixons un seuil de trafic pour diviser la station de base en deux groupes: actifs et désactifs, nous résolvons le problème de la définition d'initialisation. Nous appliquons l'algorithme à chaque période de temps où les arêtes sont pondérées par une fonction utilitaire qui quantifie la quantité d'énergie économisée et la qualité de service en échangeant les deux sommets. Mais il semble que l'algorithme met beaucoup de limitations à ce processus, à savoir:

- Comment définir le seuil de trafic pour avoir une bonne initialisation.
- En cas d'utilisation d'un seuil, il peut sembler avantageux de désactiver une station de base sous-utilisée au cours de cette période. En termes d'économie d'énergie par rapport au niveau

du système, les utilisateurs locaux pourraient toutefois être pénalisés si leur station de base était désactivée.

- Échanger des sommets est contre-intuitif et pourrait même s'avérer futile dans le cas de commutation des stations de bases sous-utilisées et surexploitées.
- Le temps d'exécution de l'algorithme est, comme indiqué précédemment, $O(n^3)$ et le nombre de stations de base en cours de lecture est plutôt important dans un environnement HetNet malgré le partitionnement. Par conséquent, il est impossible de réappliquer l'algorithme à chaque intervalle de temps, quand aucun changement n'a été établi sur notre réseau. Ce qui soulève la nécessité d'optimiser le processus de mise à jour des deux groupes.

En conséquence, nous examinons la version étendue et révisée proposée par Fidducia et Mattheyses, car elle résout le problème de l'échange de deux sommets et permet la gestion de partitions de taille non équilibrée et, avec l'utilisation de la structure de données pour stocker les gains, la complexité

temporelle est réduite. Nous envisageons également le clustering, qui divise le réseau en petits réseaux faciles à gérer. Cela nous permet d'élaborer un algorithme meilleur et plus facile à gérer.

Algorithm 1: Kernighan-Lin

Input: Graph $G=(V,E)$
Output: V_1 and V_2
Data: Cout Interne I , Cout Externe E , D_{max} , poids w

```

1  $D_{max} \leftarrow \infty$ 
   /* INITIALISER  $V_1$  et  $V_2$  */
   /* si la partition  $V_1$  contient des noeuds débloqués, on effectue les
      tentatives d'échange */
2 while  $DEVEROUILLE(V_1) \neq TRUE$  do
3   for  $a \in V_1$  do
4     if  $DEVEROUILLE(a) \neq TRUE$  then
5       for  $b \in V_2$  do
6         if  $DEVEROUILLE(b) \neq TRUE$  then
7            $D(a) = E_a - I_a$   $D(b) = E_b - I_b$   $g(a, b) = D(a) + D(b) - 2 * w_{ab}$ 
           Table( $g(a, b)$ )
8          $D_{max} = MAX(g(a, b))$ 
9         MAJ( $V_1 - \{a\}$  and  $V_2 - \{b\}$ )
10        if  $D_{max} > 0$  then
11           $D_{max} = D(a) + D(b) - 2 * w_{ab}$ 
12          SWAP(a,b)
13          VERROUILLER(a,b)

```

3.3.2 Fedducia-Mathesis

Comme indiqué dans la section précédente, *Kernighan-Lin* présente de nombreuses limitations (algorithme lent, conçu pour des partitions équilibrées, attribution initiale aléatoire), Fidducia et Mattheyses ont présenté une version modifiée de *Kernighan-Lin* qui minimise le coût de coupe. Cette méthode est basée sur les filets plutôt que sur les bords. La méthode *FM* effectue des itérations au cours desquelles un seul sommet est déplacé au maximum une fois. Par conséquent, l'algorithme

permet de gérer les partitions non équilibrées. En outre, la bisection correspondant à la plus petite taille de découpe est utilisée comme entrée pour la prochaine itération, ce qui diminue la complexité temporelle globale de $O(n^2)$. Cependant, le problème du partitionnement initialement fixé mais l'aspect aléatoire persiste. Ainsi que la limitation introduite par le facteur de ratio qui a été introduit pour empêcher le regroupement de tous les sommets dans une partition.

3.3.3 Spectral Clustering

Comme nous pouvons le constater, notre principale problématique avec ces algorithmes réside dans la fixation de la taille des groupes et dans la meilleure gestion de la stratégie de déchargement. Aucun des algorithmes précédents ne correspond à la description pour produire une combinaison optimale. Notre problème est défini par des contraintes, respectivement, au niveau système et local à savoir l'économie d'énergie du réseau, ainsi qu'une QoS garantie et un déchargement des stations de base optimal. Certes, que nous ne pouvons pas exprimer dans une fonction coût.

Premièrement, nous commençons par diviser le réseau en graphes, de manière à éviter au moins la taille des groupes. Parce que la mise en cluster ou la détection de communauté se réfère à la division du réseau, en fonction du modèle des bords. Par extension, les groupes formés seraient étroitement liés.

Contrairement au partitionnement graphique, la détection de communauté ne vous oblige pas à spécifier la taille des groupes à l'avance. En fait, le but est de trouver une division cohérente et naturelle, ce qui est plus utile pour analyser les modèles et comprendre la structure du réseau. De plus, avec le clustering, le réseau ne sera pas divisé si aucune bonne partition n'existe; contrairement au graphe de partitionnement, qui se contentera de la moins mauvaise division. Par conséquent, le fait de regrouper les stations de base en ensembles de clusters basés sur la similarité de distance et la dissimilarité de charge permettra une meilleure coordination entre les stations de base et une stratégie d'activation / désactivation efficace. Les caractéristiques de similarité définissent la capacité de coordination avec les stations de base voisines, tandis que la dissimilarité de la charge reflète la volonté de coordination car elle ne procure aucun avantage en termes de transfert du trafic vers une station de base avec la même charge. Ainsi, il est possible de trouver une coopération optimale entre deux BS, où nous pouvons appliquer la migration de charge et garantir les contraintes de qualité de service, ainsi qu'une réduction de la consommation d'énergie.

Cependant, pour effectuer le regroupement, il existe de nombreux mécanismes disponibles dans la littérature. Mais comme le réseau est vaste, la cartographie d'une matrice de similarité est exhaustive. Nous avons constaté que les méthodes de regroupement se divisent en deux catégories: centralisée et décentralisée. La centralisation, comme la classification k-mean, nécessite la connaissance de toute la typologie, tandis que la décentralisation dépend des informations communiquées par les voisins. En conséquence, avec le clustering décentralisé (basé sur la recherche d'égal à égal), nous allons réduire le temps et l'espace mémoire nécessaires pour cartographier la matrice de similarité. Après coup, nous devons envisager une coordination entre clusters.

Une fois les clusters formés, les stations de base en mode veille ou avec la charge de trafic la plus faible répartissent leur trafic vers les stations de base voisines au sein du cluster. En général, les méthodes de clustering peuvent être classées selon les catégories suivantes:

- Regroupement en partition,
- Regroupement hiérarchique: en créant une coupe hiérarchiques des objets selon certains critères,
- Regroupement basé sur la densité: en créant des partitions selon la densité et la connectivité des données.

Dans notre cas, nous abordons notre problématique en utilisant des méthodes de partitionnement. En effet, en utilisant cette méthode, nous obtenons un nombre k de partitions où chaque partition correspond à un cluster. Les clusters sont formés pour optimiser un critère de partitionnement d'objectif [14].

Afin d'établir un cluster significatif, nous identifions les similitudes entre les sommets, que nous noterons $(s_{i,j})_{i,j}$, pour les regrouper en conséquence. Ainsi, nous regrouperons des BS ayant des caractéristiques similaires, de manière à ce que les éléments les plus similaires soient identiques, alors que ceux qui sont les plus dissemblables sont placés dans d'autres graphes. Par conséquent, nous produisons un cluster compact. Comme nous l'avons mentionné précédemment, les caractéristiques de similarité utilisées sont les emplacements des stations de base et leurs charges pour calculer ces similitudes, nous avons constaté qu'il existe plusieurs méthodes dont trois sont répertoriées ci-dessous:

- *voisinage* – ϵ : qui connecte tous les sommets dont le poids par paire est inférieur à ϵ .
- *k – plus proche voisin*: qui regroupe les sommets en fonction des k-plus proches voisins. Signifiant, v_i, v_j sont connectés si v_j fait partie des k-voisins de v_i
- *Voisinage entièrement connecté*: il connecte tous les sommets avec une arête et le pondère selon une mesure de similarité tel que le noyau Gaussien ou le format RBF, qui est couramment utilisé pour la localisation dans l'apprentissage machine pour mesurer la similarité entre objets. Il est défini comme suit:

$$s_{ij} = \exp \frac{-\|x_i - x_j\|^2}{2\sigma^2} \quad (3.1)$$

Où σ contrôle la largeur du voisinage. Ce qui joue en fait le même rôle que *voisinage* – ϵ .

1. *Similarité basée sur la distance*:
2. *Dissemblance basée sur la charge*: après avoir regroupé les stations de base en fonction de leurs distances, nous ajoutons une deuxième couche de formation dans laquelle nous regrouperons les stations de base voisines au terme de leur mandat en charge. Supposons que les deux stations de base ont respectivement des charges ρ_i et ρ_j . Naturellement, nous devons regrouper les stations de base avec des charges différentes afin d'appliquer la migration du trafic. En fait, si $\rho_i \gg \rho_j$, la BS i est en mesure de prendre en charge le trafic de la BS j . Par conséquent, plus la différence de charge est grande, plus il est probable que les stations de base soient regroupées. Ainsi, sur la base de ces observations, nous définissons la matrice de dissimilarité à l'aide de la similarité gaussienne.

$$s_{ij}^l = \exp \frac{\|\rho_i - \rho_j\|^2}{2\sigma_l^2} \quad (3.2)$$

Où σ_l contrôle la largeur de la similarité. Pour combiner ces deux métriques, conformément à [11], afin de préserver les caractéristiques des similitudes Gaussiennes, il est préférable de joindre les metrics de manière non linéaire mais exponentielle:

$$S_{ij} = (s_{ij}^d)^\theta * (s_{ij}^l)^{(1-\theta)} \quad (3.3)$$

où $0 \leq \theta \leq 1$ contrôlent l'impact de chaque mesure, dans notre cas puisque nous visons un impact égal, nous définissons $\theta = 1/2$.

Après avoir défini notre matrice de similarité, nous effectuons un clustering BS. En fait, il existe de nombreux mécanismes de clustering disponibles. Dans les sections suivantes, nous examinerions trois des nombreuses méthodes de partitionnement. Le but de cette étude est d'analyser les différentes méthodes de classification que nous pouvons utiliser.

Le regroupement spectral s'appuie sur l'algèbre linéaire pour analyser les matrices d'adjacence d'un graphique [22], le filtrage spectral est une application de la théorie spectrale qui utilise un graphe Laplacien et la matrice de similarité, que nous avons introduits précédemment, pour construire des clusters. Les sommets de la même partition sont étroitement liés alors que nous minimisons les sauts entre les clusters. Le processus de clustering spectral est le suivant:

- Construire la matrice de similarité comme présenté à la section précédente.
- Définir la matrice L de Laplacien
- Intégrer les sommets en utilisant les valeurs et les vecteurs propres de la matrice Laplacienne L .
- Appliquer l'algorithme classique du clustering (comme k-means) sur les valeurs propres afin de déduire les clusters.

Pour construire la matrice Laplacienne:

$$L = D - S \tag{3.4}$$

où $D = (d_{i,j})_{i,j}$ est la matrice degré diagonale $d_i = \sum_j s_{i,j}$.

Le nombre de clusters k est étroitement lié aux valeurs propres de L . En fait, dans le cas où nous n'avons pas obtenu de valeurs propres nulles, nous pouvons choisir k tel que $\lambda_1, \dots, \lambda_k$ sont négligeable par rapport à λ_{k+1} , c'est-à-dire:

$$k = \operatorname{argmax}_i (|\lambda_{i+1} - \lambda_i|) \tag{3.5}$$

En réalité, la propriété 3.5 est un résultat direct de la théorie de la perturbation; si nous observons le cas idéal de k-graphes complètement déconnectées, les valeurs propres nulles ont une multiplicité de k et l'écart entre λ_{k+1} et λ_k serait plutôt très grand. À cette fin, une distribution uniforme de sommets

se traduirait par un intervalle uniforme entre les valeurs propres consécutives. En conséquence, plus notre graphe est grand, plus le nombre k est grand. En conclusion, l'avantage de la classification spectrale est qu'on nécessite pas la forme des graphes ni le nombre de graphes au préalable. Et en utilisant le graphe de similarité, nous pouvons créer des clusters significatifs. À cet effet, le regroupement de nos stations de base facilite la stratégie d'activation/désactivation, car les clusters peuvent remettre efficacement les UE, desservies par une station de base légèrement chargée à désactiver, à une station de base active. Ainsi, le nombre de BS désactivées peut augmenter.

3.3.4 La recherche Taboue

La recherche taboue est l'une des approches heuristiques pour la réalisation de problèmes combinatoires. En effet, un bon nombre d'algorithmes heuristiques sont des méthodes appelées 'Hill Climbing'. Fidèlement à son nom, ces processus partent d'une solution initiale, modifient certains paramètres de la solution courante jusqu'à ce que l'algorithme n'est plus capable de provoquer de modifications possibles. A cet effet, le processus sera bloqué dans un optimum local.

Contrairement à ces approches, la recherche taboue est développée pour empêcher le processus d'être bloqué dans les optimaux locaux. Pour ce faire, cette approche permet d'appliquer des changements contrôlés par la liste d'historique.

En effet, Golver et al [19] ont développé la recherche taboue qui consistant en un ensemble de règles d'évaluation pour diriger la recherche rendant la méthode plus flexible et adépte à retrouver des solutions intéressantes. Mais aussi, c'est un ensemble des règles permettant la continuation de l'exploration de l'ensemble des solutions au delà de l'obstacle des optimum locaux [18].

De ce fait, on peut voir la recherche taboue comme une méthode d'amélioration locale traditionnelle, explorant itérativement l'ensemble des solutions d'un problème d'optimisation. En effet, l'algorithme se déplace d'une solution courante S à une nouvelle voisine $N(S)$, tout en évaluant la fonction objectif.

On définit la solution voisine $N(S)$ comme étant l'espace des solutions obtenues suite à une transformation locale de la solution S . Pour notre problématique de commutation active/désactive, une telle transformation est le changement d'état.

Toutefois, l'exploitation d'une mémoire; la notion de la liste taboue; nous amène à éviter de retourner aux solutions déjà rencontrées puisqu'elles seront déclarées comme taboues. Sauf qu'on peut risquer de rendre la recherche moins flexible si la taille de la mémoire taboue est large, puisque dans ce cas les transformations taboues seront maintenues en permanence. Par conséquent, on doit être prudent durant la définition de la durée de la liste taboue afin de permettre le recyclage des choix passés et ainsi mener une exploration plus large.

Pour l'instant, nous avons introduit les principes qui dirigent et impactent considérablement la recherche taboue, à savoir : la mémoire taboue (notée liste taboue ou T), la durée de vie taboue, et enfin la fonction objective. À l'aide de l'algorithme ci-dessous (Algorithme 2) , nous allons expliciter la méthode de recherche. Soit le graphe $G = (V, E)$, on définit la solution S comme étant l'ensemble des sous-ensembles V_1 et V_2 . Les transformations qui permettent de passer d'une solution S à une autre voisine $N(S)$ consistent à transférer un vertice v_i d'un sous ensemble à l'autre. La fonction objective utilisée correspond à une sommation du poids des vertices dont les extrémités ne sont pas dans le même ensemble:

$$Cutsizes(S) = \sum_{v \in V_1} \sum_{u \in V_2} E(v, u) \quad (3.6)$$

La liste taboue sera gérée comme une liste circulaire, en éliminant l'élément dont la durée de vie est expirée pour faire place au nouveau. Puis, on interdit les transformations inverses des éléments gardés dans la liste T.

Certes, quand la liste taboue interdit l'inversement d'une solution, jusqu'à après un certain nombre d'itérations, elle interdit aussi tout ensemble de solutions possiblement pas encore visitées. Ainsi, un mécanisme de revocation est introduit si une solution est jugée intéressante. C'est le concept de critère d'aspiration. Celui utilisé, le plus simple, consiste à révoquer le statut tabou

à une transformation si elle induit une solution de qualité plus supérieure à celle présente[21].

Algorithm 2: Tabu Search

Input: Graph $G=(V,E)$

Output: $S=(V_1,V_2)$

Data: BestSolution, longueur de la liste Tabu size l, Liste Tabu T, $noimprov_{max}$

/ INITIALISER $S_0 = (V_1, V_2)$ */*

- 1 $BestSolution \leftarrow Cutsizes(S_0)$
- 2 $BestCandidate \leftarrow S_0$
- 3 **while** ($noimprov \leq noimprov_{max}$) **do**
- 4 **for** $a \in V_1$ **do**
- 5 $Cost \leftarrow Cutsizes(V_1 - \{a\}, V_2 \cup \{a\})$
- 6 **for** $b \in V_2$ **do**
- 7 $Cost \leftarrow Cutsizes(V_1 \cup \{b\}, V_2 - \{b\})$
- 8 $S \leftarrow MIN(Cost)$
- 9 $node \leftarrow MIN(Cost)$
- 10 **if** $node \notin T$ **then**
- 11 $BestCandidate \leftarrow S$
- 12 $BestSolution \leftarrow Cutsizes(S)$
- 13 $T \leftarrow node$
- 14 MAJ(V_1 and V_2)

3.4 Conclusion

Dans ce chapitre, nous avons décrit les différents algorithmes présents dans la littérature. Dans le chapitre suivant, nous expliquons l'environnement du simulateur utilisé et notre stratégie de déploiement de commutation des BS. Une analyse des résultats obtenus est également présentée à la fin du chapitre V.

Chapitre 4

APPROCHE PROPOSÉE DE GESTION DE CONSOMMATION D'ÉNERGIE

4.1 Introduction

Dans ce chapitre, nous présentons notre proposition et le travail que nous avons réussi à mettre en œuvre.

Dans un premier temps, nous décrivons le simulateur LTE qui définit la structure de notre travail, nous élaborerons sur les modèles de canal établis, la planification de l'association des utilisateurs dans notre simulateur ainsi que les paramètres utilisés pour nos simulations. De plus, nous décrivons notre processus de mise en œuvre. Enfin, nous présentons nos résultats et analysons les performances de notre algorithme.

4.2 Description du Simulateur

Pour notre simulateur, nous avons opté pour un simulateur basé sur le LTE, car notre préoccupation est principalement basée sur le signal de liaison descendante et sur le fait que la 5G utilise

le même signal que le LTE. Rétrospectivement, notre trame de signal serait une OFDM avec 7 symboles et d'un temps de slot TS de 1 ms .

De plus, notre réseau étant hétérogène, nous déployons des stations de base de petite et de grande portée afin d'améliorer la qualité de service des utilisateurs. Ainsi, notre simulateur devrait prendre en compte les deux environnements (macro et picocellule).

Pour mettre en œuvre et déployer notre réseau, nous allons simuler notre réseau local hetnet sur deux niveaux: Simulation de niveau de liaison (LLS) qui concerne le canal radio entre l'émetteur d'antenne et le transmetteur d'antenne où nous générons un modèle de canal qui couvre la propagation des macro et micro-cellules urbaines ainsi que des simulations au niveau du système qui représentent les fonctionnalités entre BS et UE. Dans cette couche, nous décrivons l'allocation des ressources, la planification de session et l'association d'utilisateurs ce qui nous fournirait suffisamment de données (tels que le SINR et le débit) pour évaluer les performances de notre réseau sur une longue période.

Ceci dit, notre simulateur nous fournira des indicateurs de performance clés dont nous avons besoin pour notre algorithme d'efficacité énergétique.

4.2.1 Structure du Simulateur

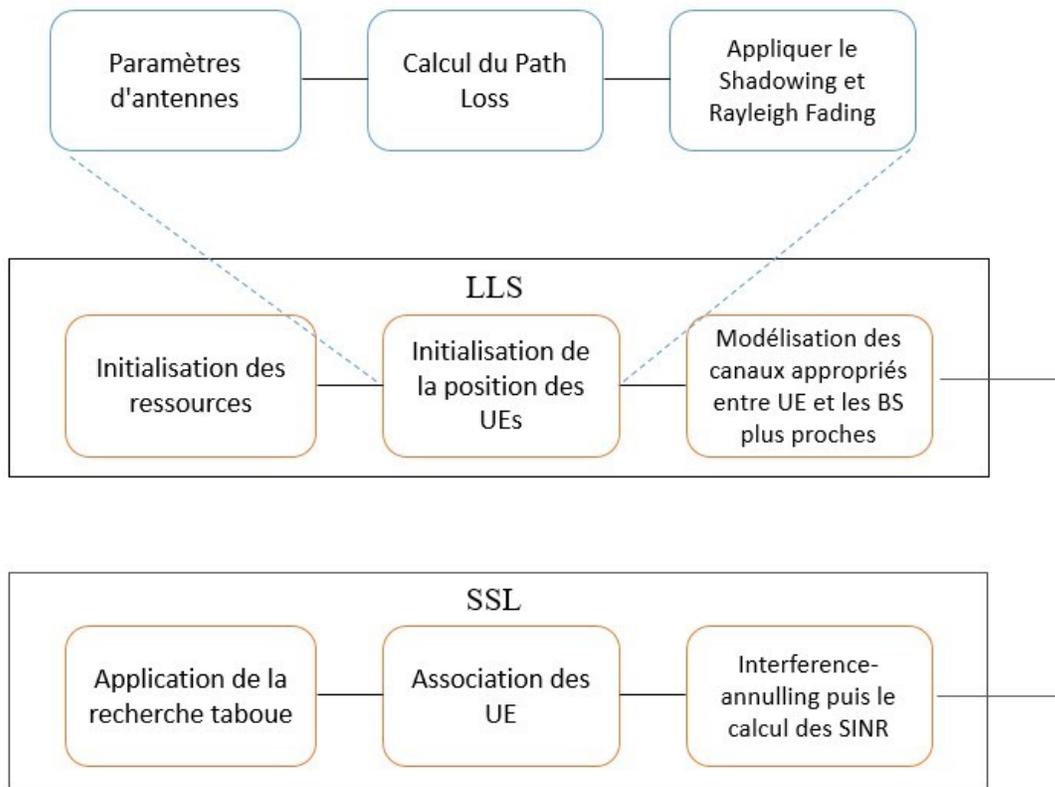


Figure 4.1 – Structure du simulateur

La figure 4.1 représente un graphique décrivant l'enchaînement du processus de construction de notre simulateur.

Nos stations de base macro sont situées dans les centres des cellules outdoors hexagonales formant une grille où les cellules pico de couverture circulaire seront dispersées autour de chaque station. Les utilisateurs mobiles sont dispersés aléatoirement autour de la station macro, en changeant leurs positions chaque TTI avec une vitesse v .

Puisque nos utilisateurs sont mobiles, la mesure de la qualité du canal entre les stations de base environnantes et les utilisateurs nous pousse à générer un canal à chaque TTI. L'efficacité de nos résultats d'algorithme d'activation ne devrait pas refléter uniquement sur le modèle de réseau pendant un TTI, mais plutôt sur une longue période de temps telle que celle d'une heure. En rétrospective, nous allons calculer la qualité du canal pour différents scénarios TTI d'une durée de

plus d'une heure et faire la moyenne de ces résultats dans le but d'avoir une moyenne statistique du canal.

Modèle du canal:

La définition du modèle de canal comprend deux type de fading: large scale fading (LSF) et small scale fading (SSF). LSF se produit lorsqu'un obstacle se positionne entre le récepteur et l'émetteur. Cependant, la puissance du signal reconnaît une fluctuation énorme à cause de deux composants du LSF: pathloss et shadowing. Quand au Fading à petite échelle, le signal reçu est soumis à des fluctuations rapides pendant une durée et une distance courtes.

D'une part, notre pathloss suivra le modèle de propagation COST Hata qui est bien reconnu pour les scénarios de propagation urban des systèmes fonctionnant dans des gammes de fréquences jusqu'à 2GHz. Par ailleurs, l'effet pathloss et celui du shadowing sont des modèles qui dépendent de la position de l'UE qui varie dans le temps. Par contre, le fading à petite échelle est un modèle qui dépend que du temps.

Dans notre version du simulateur, nous allons nous baser sur le modèle du canal Rayleigh pour un mode de transmission à deux antennes, et dont on génère des coefficients de canal pour chaque lien UE-BS. Ceci dit, la complexité de génération des coefficients de canal pour un seul scénario, varie exponentiellement par rapport au nombre total des liens entre un utilisateur et l'ensemble des antennes. Par conséquent, pour réduire la complexité des calculs, nous allons pré-générer uniquement les coefficients de canal entre un utilisateur et les antennes des N plus proches stations de bases. À fur et a mesure, nous générons des liens à la demande.

Si nous voulons accorder plus de précisions aux effets du fading à petite échelle, d'une manière plus rapide et moins complexe, nous pouvons générer en mode offline les coefficients canaux pour des différents scénarios de délai et d'angle d'arrivée . Ceci dit, la moyenne statistique de ces réalisations durant toute une période d'une heure nous donnera une mesure statistique plus réaliste de l'état du canal pendant cette période.

L'algorithme recherche Tabou

Une fois le réseau déployé, l'association d'utilisateurs configurés, et les canaux et les éventuels trafics générés; Nous allons configurer la matrice de contiguïté pour tracer un graphique pondéré représentant notre réseau. En fait, la matrice d'adjacence est une matrice symétrique avec des zéros sur sa diagonale. En outre, le relation de contiguïté serait basée sur le voisinage: ce qui signifie que deux stations de base b_1 et b_2 sont considérées comme voisins si la distance euclidienne entre eux, d_{b_1,b_2} , est inférieure que le rayon de couverture ϵ_d :

$$E = \{e_{bb'} = 1; d_{b_1,b_2} \leq \epsilon_d\} .$$

Considérons maintenant le graphe $G=(B,E)$, où B est l'ensemble des stations de base et E l'ensemble des arêtes indiquées par la matrice d'adjacence. En outre, chaque lien entre deux BS sera attribué un poids, w , représentant le degré de similarité ou de dissimilarité entre eux. Pour rappel, la décision d'activation / désactivation prend en compte à la fois la charge de trafic d'une BS et la qualité de service de l'UE. En outre, pour assurer l'efficacité énergétique et le débit du réseau, nous devons simultanément restreindre l'interférence entre les cellules. Dans ce contexte, nous avons conçu une métrique eq.4.1:

$$M = \left(\frac{P_{Totale}^0 - P_{Totale}^{ON}}{P_{Totale}^0} \right)^{w_p} \left(\frac{T_{ON}}{T_0} \right)^{1-w_p} \quad (4.1)$$

Cette fonction est définie comme le produit du gain en réduction de puissance et le gain en débit. $P_{Totale}^0, P_{Totale}^{ON}$ représente respectivement la puissance totale consommée par le réseau et la puissance totale consommée par le groupe des station de base actives V_{ON} , eq.2.5. Dans notre étude, nous visons à augmenter cette metrique, tandis que nous diminuons le coût de la coupe entre l'ensemble des BS actives et les BS désactivées, $cutsiz(e(V_{ON}, V_{OFF}))$, au niveau du partitionnement des graphes. Par conséquent, nous définissons le coût de la coupe selon l'équation (4.2):

$$cutsiz(e(V_{ON}, V_{OFF})) = 1 - M = 1 - \left(\left(\frac{P_{Totale}^0 - P_{Totale}^{ON}}{P_{Totale}^0} \right)^{w_p} \left(\frac{T_{ON}}{T_0} \right)^{1-w_p} \right) \quad (4.2)$$

La consommation totale du réseau est défini autant que sommation de la puissance de consommation totale des stations de base macro-cellule et pico-cellule, à savoir l'équation 2.5 .

$$P_{Totale}^0 = \sum_{b_1}^{Macro} (P_{macro}^{RF} * N_{macro}^{antenna} + \frac{P_{macro}^{tx} * \rho_{b_1}}{\eta_{macro}^{PA}} + P_{macro}^{static}) + \sum_{b_2}^{Pico} (P_{pico}^{RF} * N_{pico}^{antenna} + \frac{P_{pico}^{tx} * \rho_{b_2}}{\eta_{pico}^{PA}} + P_{pico}^{static}) \quad (4.3)$$

$$P_{Totale}^{ON} = \sum_{b_1}^{V_{ON}} (P_{macro}^{RF} * N_{macro}^{antenna} + \frac{P_{macro}^{tx} * \rho_{b_1}}{\eta_{macro}^{PA}} + P_{macro}^{static}) * a_{b_1} + (P_{pico}^{RF} * N_{pico}^{antenna} + \frac{P_{pico}^{tx} * \rho_j}{\eta_{pico}^{PA}} + P_{pico}^{static}) * (1 - a_{b_1}) \quad (4.4)$$

où P_{Totale}^{ON} est la puissance totale consommée par les stations de base actives , $a_{b_1} = 1$ si b_1 est une macrocellule et $a_{b_1} = 0$ si c'est une picocellule. Supposons que K_{b_1} est l'ensemble des utilisateurs desservis par une station de base b_1 , on définit la charge de cette station de base comme suit:

$$\rho_{b_1} = \sum_{k \in K_{b_1}} \frac{\lambda_k}{r_{b_1,k}} \quad (4.5)$$

où λ_k est le débit requis par l'utilisateur k et $r_{b_1,k}$ le débit reçu défini par eq.2.7.

Tandis que T_{ON} et T_0 représentent, respectivement, le débit moyen fourni par les stations actives durant une itération de la recherche taboue et le débit moyen initial fourni par le réseau avant d'appliquer notre recherche.

Dans le problème d'optimisation, l'algorithme vise à trouver une partition avec le plus petit coût de coupe entre les bi-partitions. Dans la théorie de graphe, la fonction objective est la somme des arêtes reliant les partitions. Les algorithmes de partitionnement de graphes sont conçus pour trouver la bi-section de sorte que son coût ou son poids ait une valeur minimale par rapport à toute autre coupe possible. Par analogie avec notre travail, notre algorithme vise à retrouver la transformation (active/désactive) qui minimiserait le coût formulé par l'équation 4.2. À cet égard, l'objectif est de trouver une partition qui regroupe le plus de stations de base bien utilisées et minimise incidemment la somme de la consommation d'énergie tout en maintenant la qualité de service. Ainsi, l'équation 4.5 quantifie la quantité des ressources requises par la station de base pour atteindre le débit requis pour chaque utilisateur.

Comme nous avons expliqué dans le chapitre précédent, la recherche taboue est une technique d'optimisation. Néanmoins, sa simplicité et son applicabilité sur des bi-partitions inégales sont la

raison pour laquelle nous avons choisi de l'adapter à notre approche. Cet algorithme nécessite de commencer par une bisection initiale, puis de rechercher de manière itérative une meilleure bisection. Dans notre cas, nous commençons notre partition en regroupant d'abord les stations de base non utilisées, V_{OFF} , tandis que V_{ON} regroupe les stations de base actives. Lors de chaque itération, nous examinons chaque nœud (c'est-à-dire BS) en le déplaçant d'une partition à l'autre. Ensuite, nous procédons à l'évaluation de l'impact de la transformation ⁽¹⁾ sur le réseau. Ce qui engendre la ré-association des utilisateurs, le recalcul des interférences et enfin le calcul de la consommation totale. Nous commençons par le nœud réduisant le plus la somme de la fonction objective 4.2. À cet effet, la recherche vise à désactiver la station de base ayant le plus fort impact d'interférences qui réduit le débit moyen, sans surcharger le reste des stations de base actives ainsi qu'activer la BS dormante pour soulager les BS actives surchargées. Après l'activation ou désactivation, nous gardons la station de base (BS) dans une liste taboue pendant un certain nombre d'itérations. Comme mentionné précédemment, la liste des tabous empêche les solutions précédemment visitées d'être inversées, ainsi évite que l'algorithme tombe dans un optimum local.

Après avoir initialisé la première division, nous examinons la fonction objective 4.2 de chaque mouvement, puis sélectionnons la solution obtenue en déplaçant une seule BS, qui coûterait moins cher. Dépendamment de la valeur de w_p que la recherche sera guidée. Pour des valeurs $w_p < 0.5$, l'algorithme penche vers des solutions accentuant la QoS, sinon vers des solutions accentuant le gain de réduction de puissance pour des valeurs $w_p > 0.5$. Ensuite, nous vérifions si la BS est déjà dans la liste des tabous. Dans le cas contraire, nous procédons aux inspections suivantes:

- Une station de base sera désactivée si et seulement si tous ces utilisateurs peuvent être remis à une station voisine tout en gardant leurs QoS.
- Une station de base sera activée si et seulement si son impact en terme de service surpasse ses interférences.

1. Si la transformation est de "désactiver" une BS, nous sommes obligés de respecter la contrainte suivante:

- Les utilisateurs servis par cette BS seront remis à l'une des stations de bases actives voisines. D'ailleurs, nous vérifions dans la liste de voisinage les BS rémanentes actives

1. La transformation représente l'activation ou désactivation

qui peuvent fournir le meilleur niveau de SINR. Dans les cas où la plupart de ces utilisateurs ne parviennent pas à sélectionner une station de base, nous devrions interdire ce déplacement. On a fixé le seuil sur le nombre d'utilisateurs rejetés à 2%².

2. Si la transformation est de "activer" une BS, nous sommes obligés de respecter la contrainte suivante:

- Le service de BS en question surpasse son impact d'interférence décrit par la contrainte 4.7.

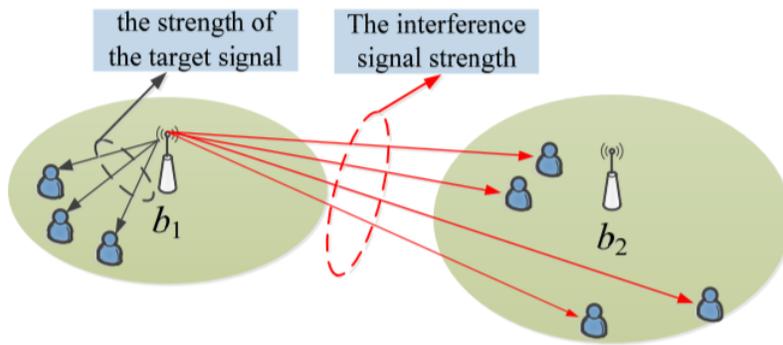


Figure 4.2 – Les puissances des signaux d'interférences et le signal reçu par les utilisateurs.

Selon la figure (4.2) tirée de [16], Γ_{b_1, b_2} est défini comme étant le rapport entre la somme des puissances des signaux reçus par les utilisateurs K_{b_1} servis par la station de base b_1 , et la somme des puissances du signal reçu de b_1 par les utilisateurs K_{b_2} servis par la BS b_2 :

$$\Gamma_{b_1, b_2} = \frac{\sum_{k \in K_{b_1}} P_{b_1}^{tx}(k)}{\sum_{q \in K_{b_2}} P_{b_2}^{tx}(q)} \quad (4.6)$$

où $P_{b_1}^{tx}$ et $P_{b_2}^{tx}$ sont respectivement la puissance reçue par les BS b_1 et b_2 . Ceci dit, la contrainte suivante calcule le taux d'interférence la station de base b_1 sur tous les BS voisines.

$$\sum_i \Gamma_{b_1, i} < threshold \quad (4.7)$$

2. Il peut être réglé selon les préférences de l'opérateur

Si l'une de ces contraintes s'applique, nous procédons au déplacement de la station de base. Sur ce, nous appliquerons toutes les modifications que ce déplacement induirait (c.-à-d. Recalculer le SINR, recalculer l'interférence et éventuellement re-associer de nouveau les utilisateurs).

Pendant chaque itération, nous gardons une trace de la fonction objective (4.2). Par conséquent, si une décision d'activer ou désactiver une BS résultant de la plus basse valeur enregistrée, la combinaison actuelle sera enregistrée comme la meilleure bisection actuelle. En fait, l'objectif principal de la liste est de limiter le sens de la recherche. Cela empêche l'algorithme de revenir à un état précédemment atteint tout en évitant d'être piégés dans un optimum local au biais de la longueur de liste. Si l'algorithme ne parvient pas à trouver une meilleure combinaison après plusieurs itérations consécutives, nous abandonnons la recherche. Pour renforcer encore nos précautions concernant l'algorithme bloqué autour d'un optimum local, nous avons mis en place deux règles:

1. Au cas où toutes les solutions voisines n'apporte à aucune amélioration du coût, nous incrémentons la variable *noimprov* puis nous forçons la transformation qui engendra la moindre augmentation du coût. Une fois la variable *noimprov* atteint un certain nombre *noimprovmax*, nous abandonnons la recherche.
2. Un nœud n'est conservé dans la liste taboue que pendant un certain nombre d'itérations.

De plus, notre algorithme atteint sa fermeture après que la variable sans amélioration ait atteint sa valeur maximale. Ces actions diversifient la recherche et permettent de sortir d'un optimum local, ce qui rend la qualité des résultats aussi convaincante. La figure ci-dessous illustre ce que nous avons décrit auparavant. Nos variables d'entrée sont: La liste Tabou T , la partition initiale V_{ON} et V_{OFF} , $noimprov = 0$, $noimprov_{max}$. La figure 4.3 , illustre ce que notre algorithme fait:

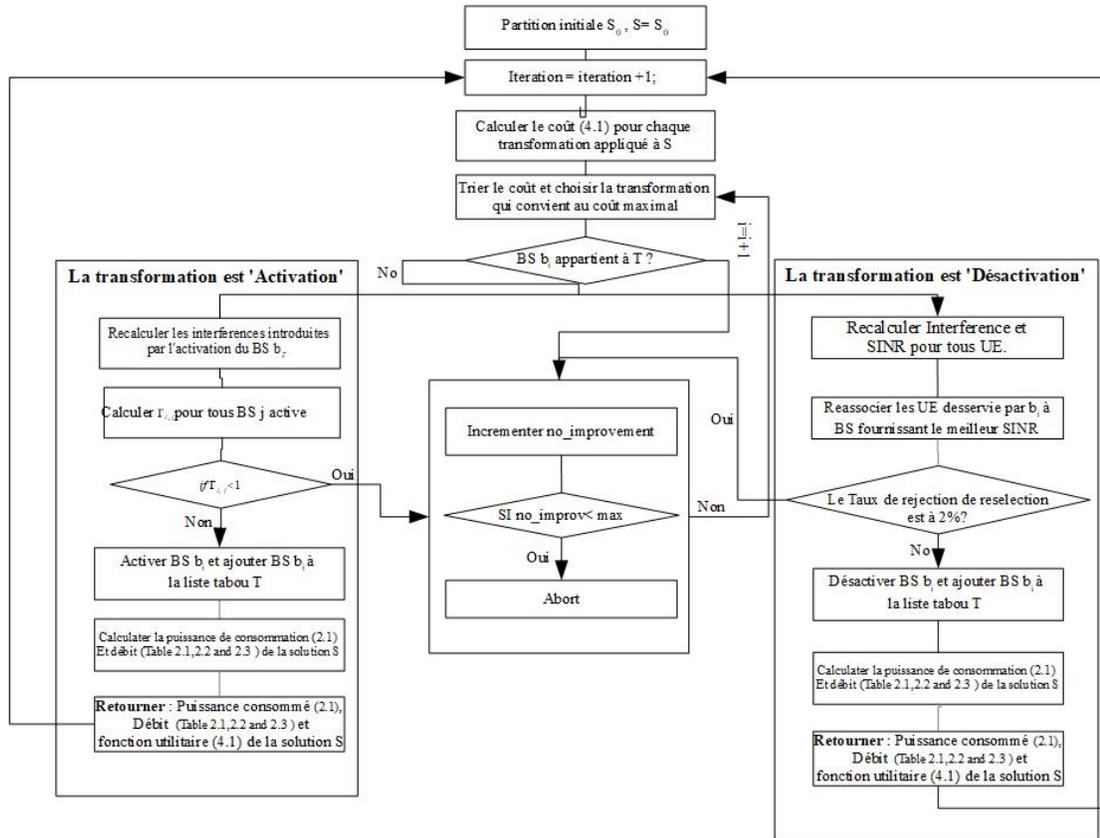


Figure 4.3 – L’algorithme Tabou

4.2.2 Paramètres de simulation

Comme décrit précédemment, nous avons utilisé un simulateur LTE générant des cellules macros positionnées aléatoirement sur site, puis il distribue de manière aléatoire les sites pico et les utilisateurs à l’intérieur de chaque région de couverture macro. Au cours de la simulation, chaque utilisateur peut être attaché à la cellule offrant la meilleure puissance RSRP, mais puisque l’utilisateur se déplace à chaque TTI, nous nous retrouvons dans le cas où l’utilisateur demandera un transfert périodiquement. Cependant, pour éviter ce cas, nous avons moyenné l’état du canal durant une heure.

Sur ce point, nous avons résumé tous les paramètres utilisés dans la simulation du déploiement du réseau dans le tableau référencé dans le travail [15].

Paramètres	Macro	Pico
Fréquence du système	2GHz	2GHz
Bande passante B	20MHz	20MHz
Nombre de Resource Blocks	100	100
Puissance de transmission P_{macro}^{tx} et P_{pico}^{tx}	43 dBm	23 dBm
Rayon de couverture R_{macro}^{cov} et R_{pico}^{cov}	1km	100m
Puissance du processeur basebande P_{macro}^{BB} et P_{pico}^{BB}	35.5 watt	2.5 watt
Puissance RF P_{macro}^{RF} et P_{pico}^{RF}	13-15 watt	2.5 watt
Efficacité d'amplification η_{PA}	23.55%	5.42%
Nombre d'antenne $N_{antenne}$	2	2
Hauteur d'antenne	30m	6m
Hauteur d'utilisateur mobile	1.5m	1.5m
Vitesse d'UE v	3Km/h	3Km/h
Taux de données requis λ	36Kbps	36Kbps
TTI	0.001s	0.001s
Temps de simulation en TTIs	3600	-
Taille du réseau N	8	15
Densité UE/cell	30,50,100,150	-

Tableau 4.1 – Tableau des paramètres de simulation

Le Tableau 4.2 résume les paramètres utilisés dans la simulation du partitionnement de graphe.

Paramètres	Valeur
Longueur de la liste taboue L	5, 6, 7, 8, 9, 10, 12, 15, 17, 20
p_{max}	100

Tableau 4.2 – Tableau des paramètres de la recherche taboue

4.3 Conclusion

Dans ce chapitre, nous avons présenté notre proposition et décrit l'environnement de simulation, ainsi que l'ensemble des paramètres de simulation. De plus, nous avons expliqué en détail l'enchaînement de l'algorithme, y compris la nature des contraintes qui conviennent à l'environnement de notre étude.

Chapitre 5

RÉSULTATS ET PERFORMANCES

Le réseau est composé de macros BS dispersées de manière aléatoire en premier, puis de pico-cellules déployées de manière aléatoire, mais centrées autour de chaque macro-cellule, et enfin, les utilisateurs sont définis de manière aléatoire autour de chaque cellule (voir la figure ci-dessous extraite de [10]). A cet effet, pour évaluer notre algorithme, nous avons décidé de l'appliquer à différents environnements.

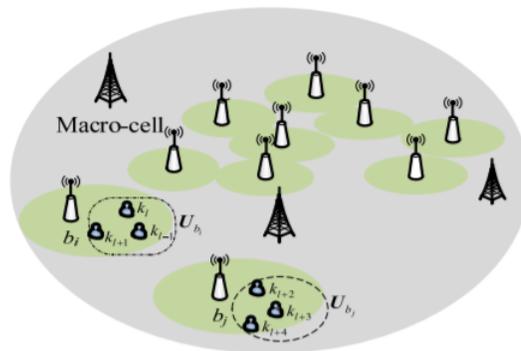


Figure 5.1 – Exemple du déploiement du réseau

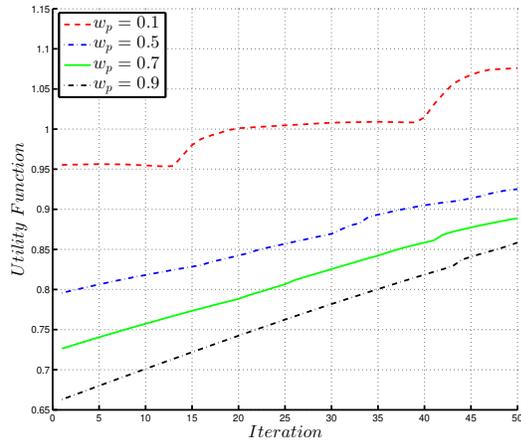
Pour l'instant, nous avons fixé le nombre de macro et de pico-cellules tout en augmentant progressivement le nombre d'utilisateurs. A priori, cette configuration permettrait d'illustrer l'impact de la quantité d'utilisateurs sur l'efficacité énergétique du réseau. Par ailleurs, pour chaque combinaison, nous devons produire plusieurs réalisations afin de conclure le Monte-Carlo; où nous déduisons une moyenne empirique des systèmes ayant plusieurs degrés de liberté.

En outre, comme nous avons expliqué dans le chapitre précédent, la recherche taboue est une technique d'optimisation combinatoire qui ne suit pas la méthode "Hill Climbing". Par conséquent, les itérations sont considérées comme solutions indépendantes, et le comportement du processus peut sembler, dans son ensemble, erratique et imprévisible. Rétrospectivement, nous produisons plusieurs résultats afin de déduire à quel moment la recherche taboue converge vers une solution optimale. Nous illustrons également la convergence de l'algorithme affectée par le nombre d'utilisateurs et la longueur de la liste taboue.

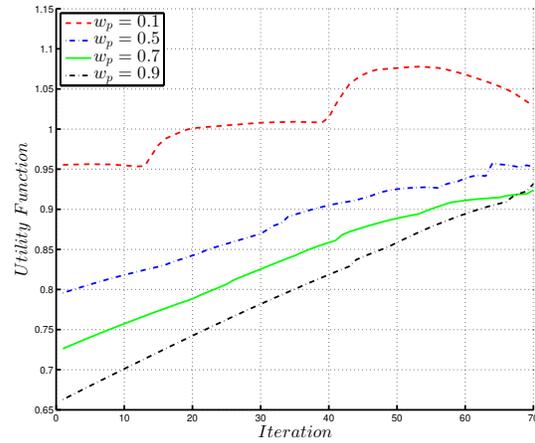
Les figures suivantes (5.5 – 5.7) décrivent les résultats obtenus pour un réseau qui déploie 8 BS macros, 5 BS picocells et 30 UEs/cellule puis 50 UEs/cellule, 100 UEs/cellule et 150 UEs/cellule.

La figure (5.2a) montre la fonction de coût (eq.4.2) pour différentes valeurs du poids de puissance w_p . Cette métrique explique à la fois la réduction en puissance et le gain en QoS. En effet, une valeur élevée de w_p signifie que la réduction de puissance a préséance sur la qualité de service. A cet égard, nous avons lancé l'algorithme tabou pour chacune des valeurs de $w_p = 0.1$, $w_p = 0.5$, $w_p = 0.7$, et $w_p = 0.9$ afin d'analyser le comportement de la recherche quand la puissance ou le débit ont préséance.

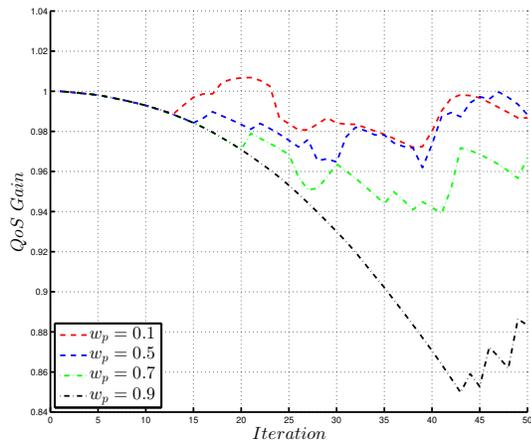
Les figures (5.2c), (5.2d), (5.2e) et (5.2f) indiquent les résultats obtenus, vis à vis des gains en QoS et en puissance, pour différents tailles de la recherche taboue. Cependant, chaque itération revient à une solution de commutation possible pour différentes valeurs de w_p et de taille de la liste taboue. Nous réalisons que pour différents w_p , la recherche ne diverge qu'après la quinzième itération, c-à-d que pendant ces itérations en dépit de notre préférence QoS ou de puissance, la recherche converge vers la même solution. Certes, les solutions varient dépendamment au poids. Cependant, selon les figures (5.2c),(5.2d), nous remarquons que les solutions retenues, quand $w_p \leq 0.5$, garantissent le QoS et même retrouvent une solution optimale qui améliore le débit de 5 à 10%. En effet, l'algorithme penche vers la commutation des stations de base sous-utilisées qui induit la diminution des interferences sur le reste des stations de base actives augmentant la QoS totale du réseau. En parallèle, le gain en puissance n'est pas meilleur que celui garantit par $w_p > 0.5$ puisque l'algorithme tend à éteindre plus de stations de base, y compris des BS non sous-utilisées, ce qui diminuent le QoS.



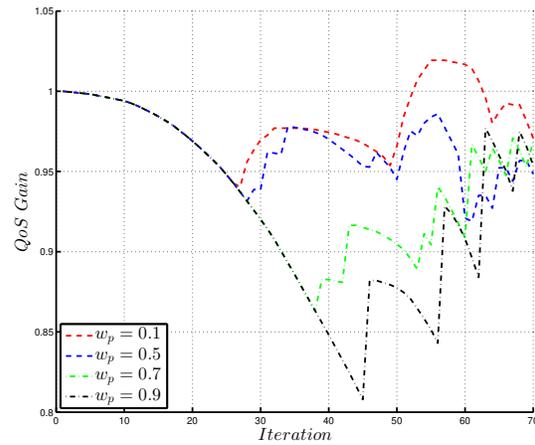
(a) Fonction utilité (4.2) établie pour plusieurs valeurs de w_p en fonction de chaque itération.



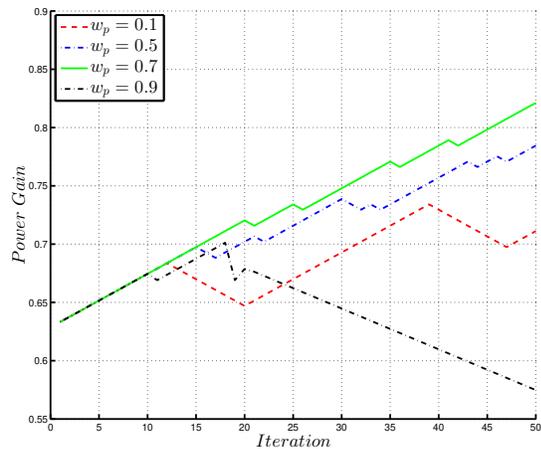
(b) Fonction utilité 4.2 établie pour plusieurs valeurs de w_p en fonction de chaque itération.



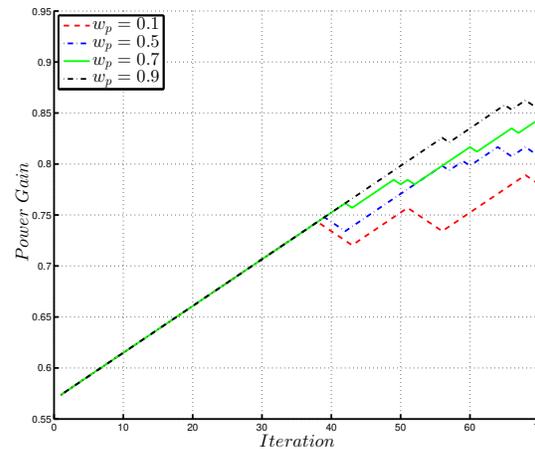
(c) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.



(d) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.



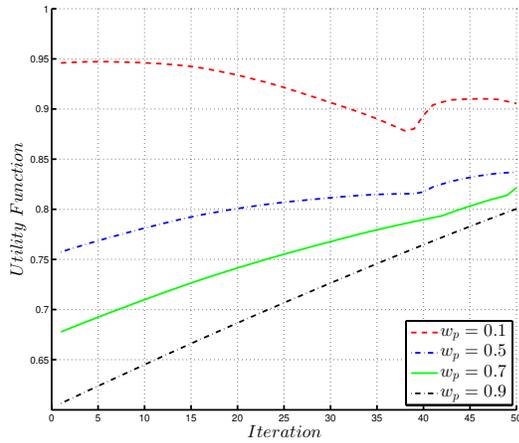
(e) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.



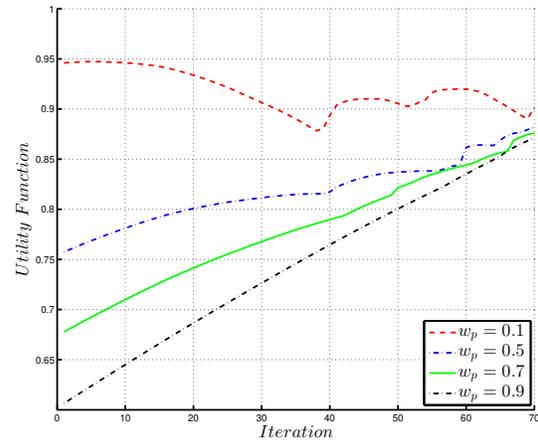
(f) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.

Figure 5.2 – Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule, et 30 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).

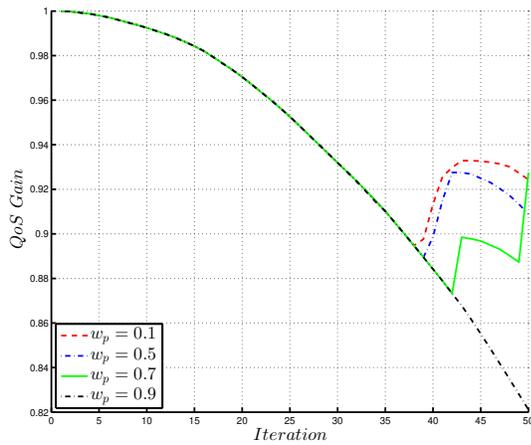
Pour un déploiement de 50 UEs/cellule, on déduit de la figure (5.3c) que notre algorithme ne retrouve plus de combinaisons qui garantissent un débit meilleur pour la taille de la liste taboue fixée à préalable (4.2). En outre, en analysant la figure (5.3d), si nous augmentons la taille de la liste taboue, nous retrouvons une solution optimale après la 50^{ème} itération. Par conséquent, nous devrions augmenter la taille de la liste au fur et à mesure que la densité du réseau augmente. Nous avons les mêmes conclusions dans le cas d'un déploiement de 100 UEs/cellule.



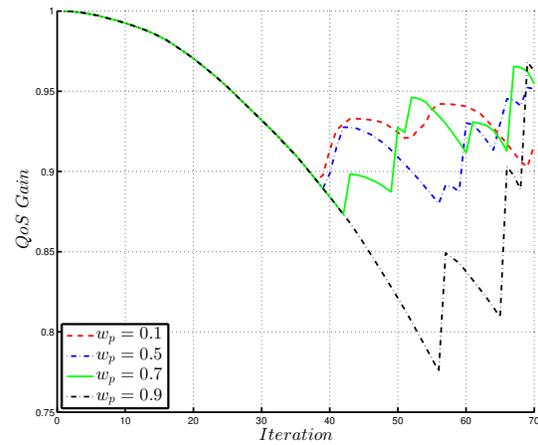
(a) Fonction utilité 4.2 établie pour plusieurs valeurs de w_p en fonction de chaque itération.



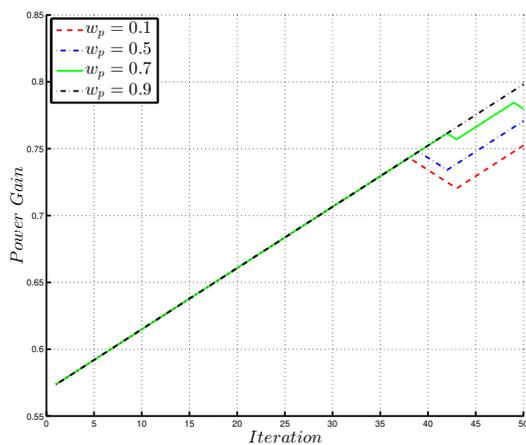
(b) Fonction utilité 4.2 établie pour plusieurs valeurs de w_p en fonction de chaque itération.



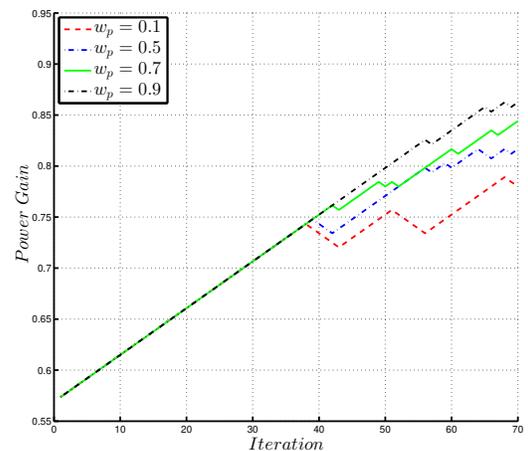
(c) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.



(d) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.

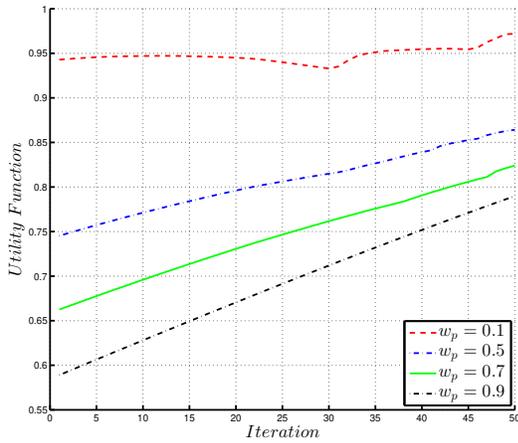


(e) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.

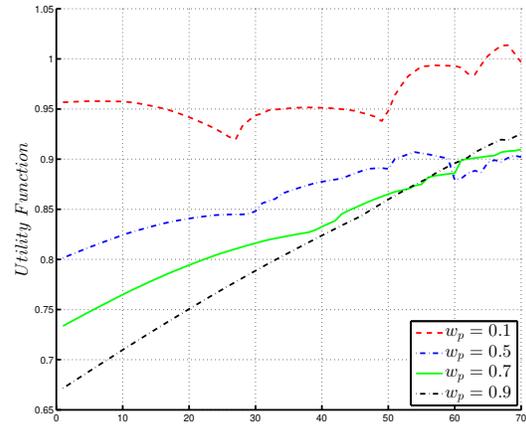


(f) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.

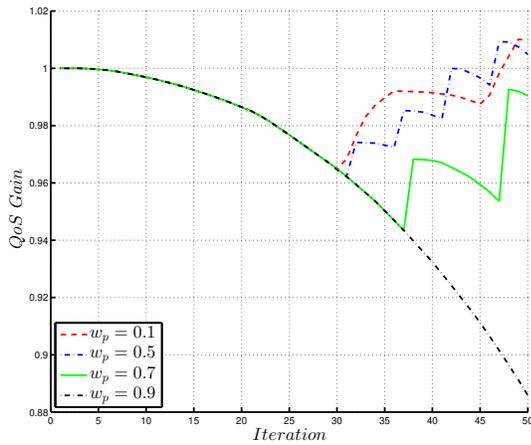
Figure 5.3 – Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule et 50 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).



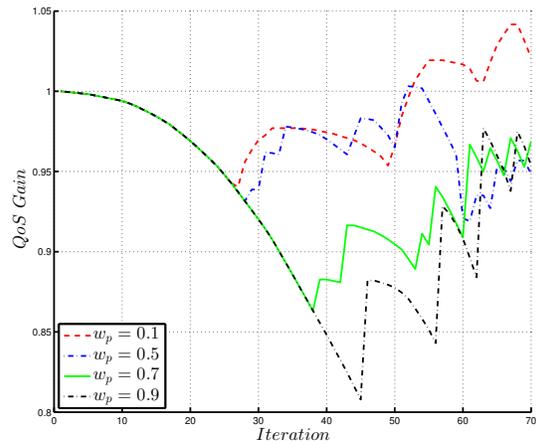
(a) Fonction utilité 4.2 établie pour plusieurs valeurs de w_p en fonction de chaque itération.



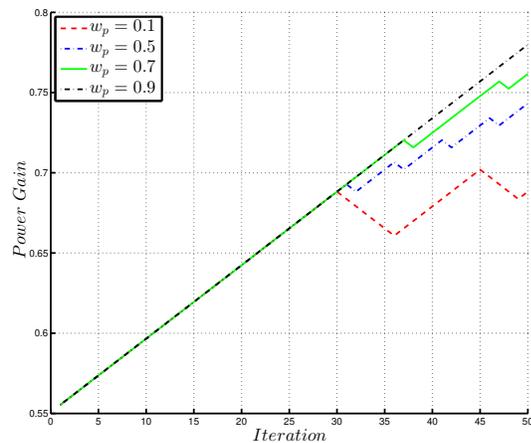
(b) Fonction d'utilité 4.2 établie pour plusieurs valeurs de w_p en fonction de chaque itération.



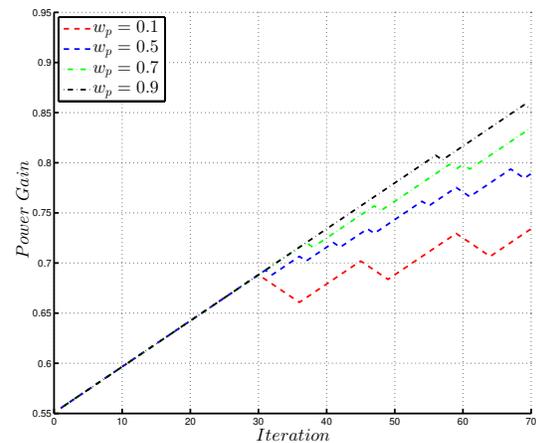
(c) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.



(d) Gain en QoS établi pour plusieurs valeurs w_p en fonction de chaque itération.



(e) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.

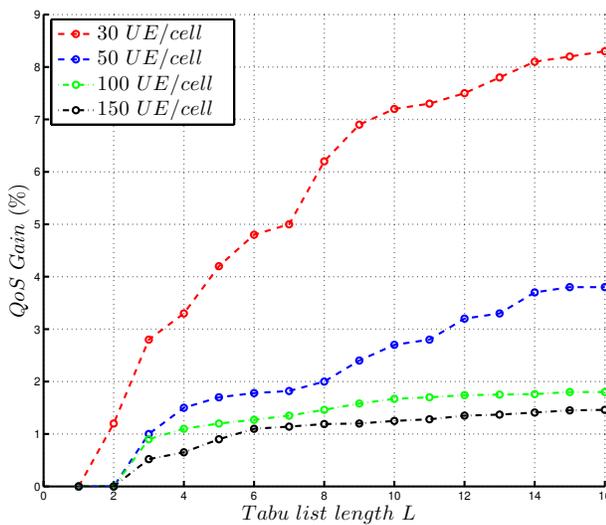


(f) Gain en puissance pour plusieurs valeurs w_p en fonction de chaque itération.

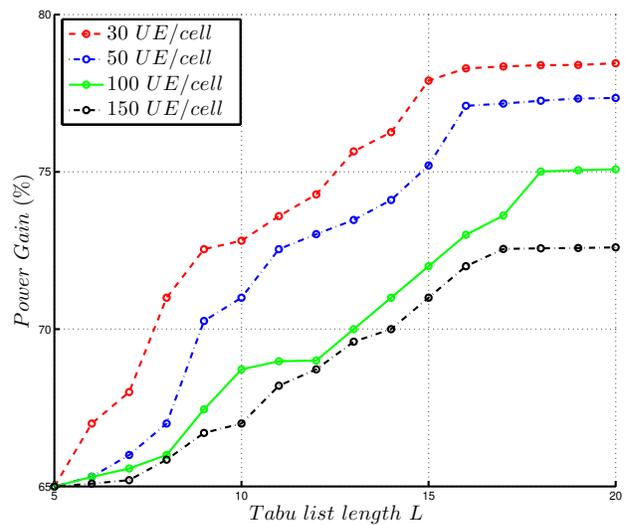
Figure 5.4 – Résultats du déploiement réseau avec 8 macro-cellules, 15 picos/cellule, et 100 UEs/cellule pour $L = 5$ dans (a), (c) et (e) et $L = 7$ dans (b), (d) et (f).

les meilleures solutions de la fonction utilité correspondent à $w_p \leq 0.5$ puisque dans ce cas on priorise le débit sachant qu'on diminue toujours la puissance totale consommée pendant la recherche. Sauf que dans ce cas, une solution optimale est plus difficile à retrouver d'où la nécessité d'augmenter la longueur de la liste taboue L . Tandis que pour $w_p > 0.5$, on priorise le gain en puissance en dépit de la qualité de service. Ainsi les résultats sont plus favorable en termes de gain en puissance qu'aux ceux de gain en QoS.

Bien que nous ayons compris l'impact du poids $w_p > 0.5$ sur les gain (en QoS et en puissance) des solutions proposées pour différentes densités d'UE, nous devons également comprendre l'impact de cette dernière sur ces gains. Ceci dit, les figures (5.5a) et (5.5b) expliquent que quand la densité en termes d'UEs/cellule augmente, l'espace des solutions améliorant la QoS devient étroit et la recherche stagne. En effet, d'après la figure (5.9), selon la longueur de la liste taboue et la densité des UEs, nous ne retrouvons plus de meilleure solution, d'où la stagnation comme illustrée dans les deux figures. En outre, lorsque la densité de la BS déployée dans une zone dépasse la densité d'UEs dans la dite zone, en référence à 30 UEs/cellule, peu nombreux pico-cellules desservent très peu d'utilisateurs. Par conséquent, notre recherche taboue permettrait de réassocier les dits utilisateurs tout en augmentant le gain de puissance et réduisant les interférences ,ainsi améliorant QoS.

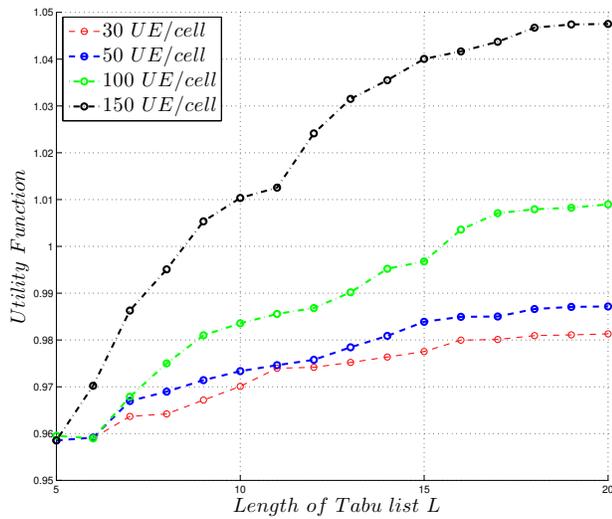


(a) Gain optimal en QoS établi pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .

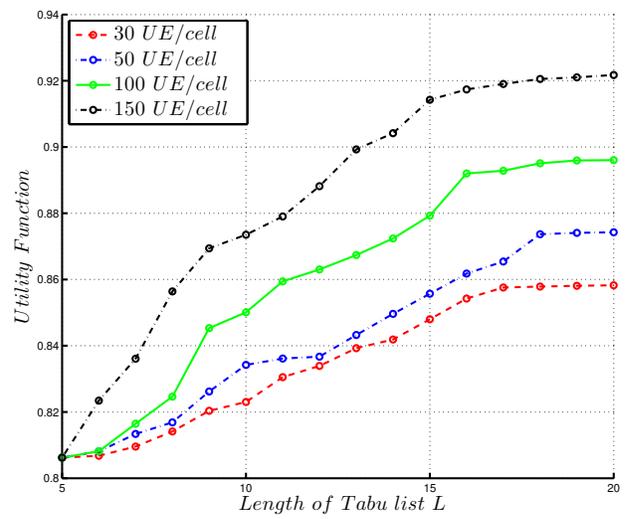


(b) Gain optimal en puissance établi pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .

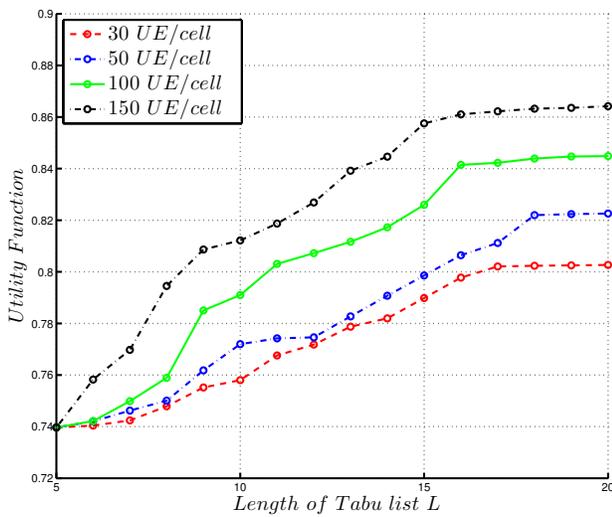
Figure 5.5 – Gains optimaux en puissance et en QoS établis pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .



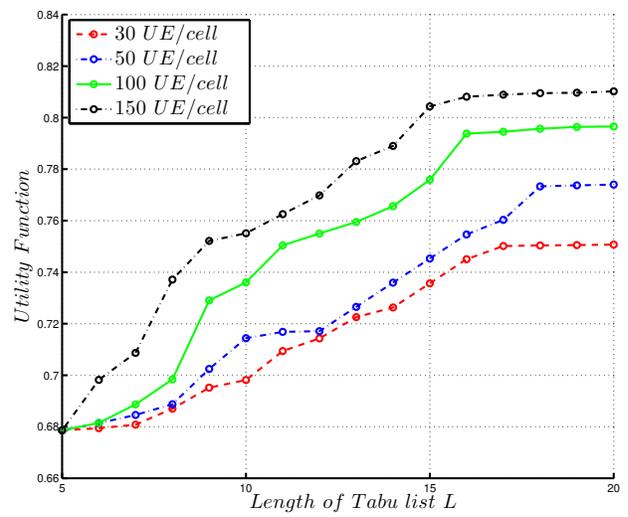
(a) Fonction utilité 4.1 $w_p = 0.1$ établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .



(b) Fonction utilité 4.1 $w_p = 0.5$ établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .



(c) Fonction utilité 4.1 $w_p = 0.7$ établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .

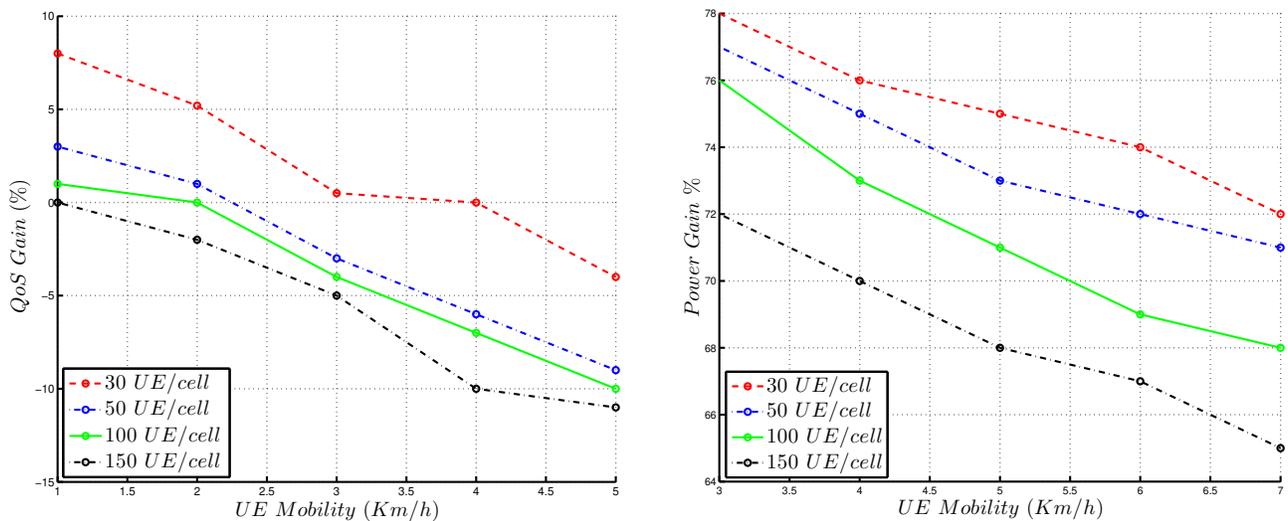


(d) Fonction utilité 4.1 $w_p = 0.9$ établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .

Figure 5.6 – Fonction utilité optimale 4.1 établie pour différentes densités en nombre d'utilisateurs par cellule en fonction de la longueur de la liste taboue L .

Le système a été testé sous trois différentes options de configurations (30, 50, or 100 UEs/cellule), dont les utilisateurs se déplaçaient à divers vitesses: 3 Km/h, 4 Km/h, 5 Km/h. En plus, le taux minimum de données requises pour les utilisateurs est supposé être 36 Kb/s 4.1. Selon [18], il a été montré qu'une valeur typique Appel VoLTE génère un trafic avec un minimum de taux de 36 kb/s.

Par conséquent, ce travail suppose que chaque utilisateur a une fréquence minimale requise de 36 kb/s.



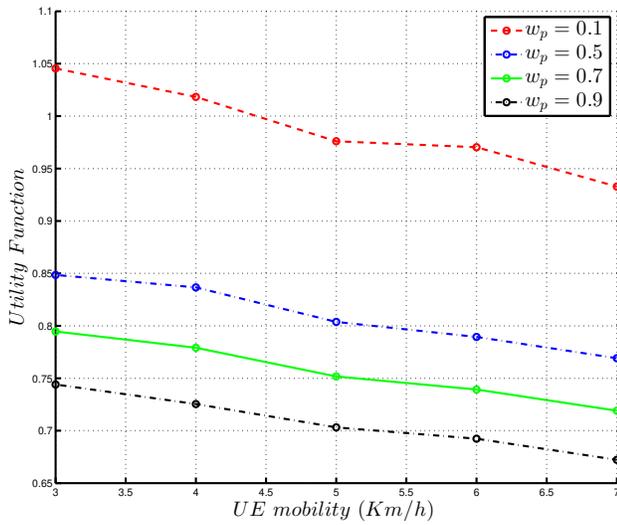
(a) Gain optimal en QoS établi pour différentes densités en nombre d'utilisateurs par cellule en fonction de la mobilité.

(b) Gain optimal en puissance établi pour différentes densités en nombre d'utilisateurs par cellule en fonction de la mobilité.

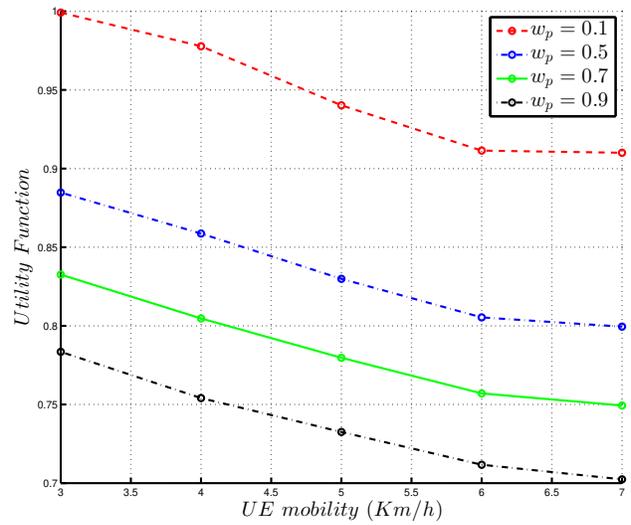
Figure 5.7 – Gains optimaux en puissance et QoS établis pour différentes densités en nombre d'utilisateurs par cellule en fonction de la mobilité.

??

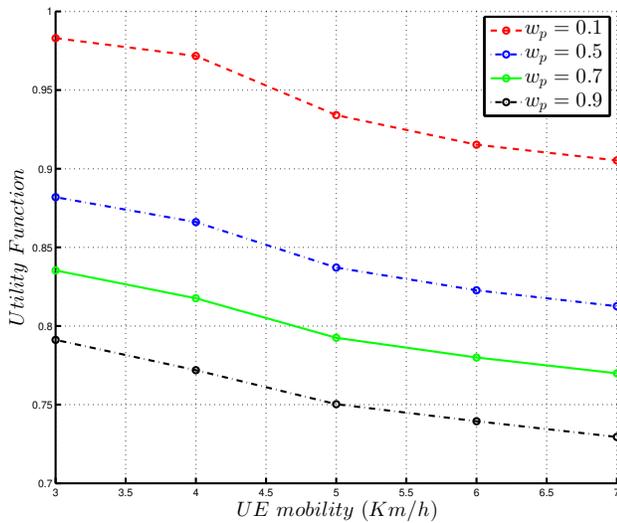
Les figures 5.7b et 5.7a illustrent, respectivement, les gains optimaux en puissance et en QoS à différentes vitesses pour différentes densités d'utilisateurs par cellule en fonction de la mobilité d'utilisateur. Dans un premier temps, on constate que le gain de puissance est une fonction croissante de la vitesse. C'est ce qui découle de la figure précédente dans laquelle la même observation a été faite et qui est illustrée par le fait qu'une quantité plus importante de ressources est nécessaire pour satisfaire les besoins et les exigences à mesure que la vitesse augmentait. Ceci est également évident pour le gain de QoS car il diminue lorsque la vitesse augmente, comme on le remarque dans la figure 5.7a. Cela peut être expliqué par le fait que des vitesses plus élevées sont synonymes des conditions médiocres du canal et donc d'un faible débit de données reçu.



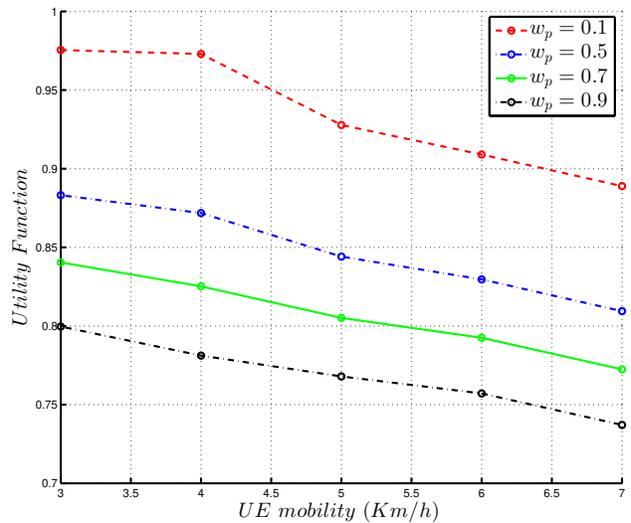
(a) Fonction utilité 4.1 30 UEs/cellule établie pour plusieurs valeurs de w_p en fonction de la mobilité.



(b) Fonction utilité 4.1 50 UEs/cellule établie pour plusieurs valeurs de w_p en fonction de la mobilité.



(c) Fonction utilité 4.1 100 UEs/cellule établie pour plusieurs valeurs de w_p en fonction de la mobilité.



(d) Fonction utilité 4.1 150 UEs/cellule établie pour plusieurs valeurs de w_p en fonction de la mobilité.

Figure 5.8 – Fonction utilité optimale 4.1 établie pour plusieurs valeurs de w_p en fonction de la mobilité.

La figure ci-dessus illustre l'impact de la mobilité des utilisateurs sur les valeurs optimales de la fonction utilité pour différents dénominateurs UE et nous suivons son comportement pour différentes valeurs de poids. Notre première observation est qu'au fur et à mesure que le nombre d'utilisateurs augmente, la valeur optimale de la fonction utilité décroît car le gain en puissance optimal ainsi que le gain en QoS diminuent, ce qui est conforme à nos conclusions précédentes (figures 5.5b et 5.5a).

Nous notons également la diminution des valeurs lorsque la mobilité augmente, comme indiqué dans les figures (5.7a et 5.7b). Enfin, les meilleures valeurs sont toujours trouvées lorsque $w_p < 0.5$.

Par ailleurs, le temps pour retrouver la solution optimale dépend de la densité des utilisateurs et de la taille de la liste taboue. Par conséquent, nous remarquons, d'après la figure 5.9, que l'augmentation de la densité des UEs par cellule nécessite l'augmentation de la longueur de la liste. En effet, quand la taille de la liste est petite, la recherche sera limitée et les solutions sont recyclées. De plus, à mesure que la densité d'utilisateurs augmente, plus de trafic est accumulé et plus de stations de base sont utilisées pour répondre à ce trafic. Cela découle de la figure 5.5b dans laquelle la même observation a été faite. Ainsi, afin d'améliorer les solutions proposées par la recherche taboue, nous devons augmenter la profondeur de la recherche, c'est-à-dire la longueur de la liste, afin d'éviter les solutions recyclées ainsi que les optimums locaux.¹

1. La moyenne haute et la moyenne basse reviennent, respectivement, au nombres d'itérations maximal et minimal pour avoir une solution optimale

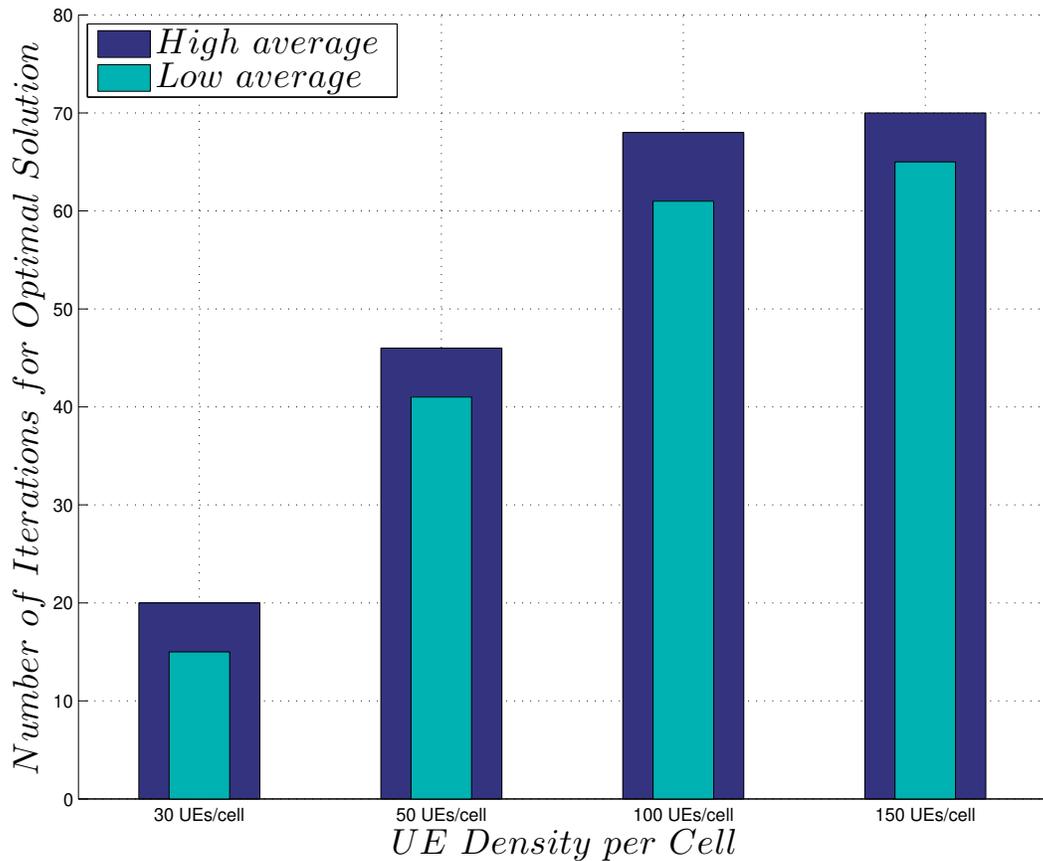


Figure 5.9 – Complexité en fonction de la densité en termes du nombre d'utilisateurs par cellule.

Nous concluons que les solutions, offrant un gain en QoS de 0-8% et un gain en puissance d'environ 78.5%, sont trouvées lorsque $w \leq 0.5$ et que, à mesure que la densité augmente, nous devrions augmenter la longueur de la liste taboue au moins à \sqrt{N} , c.à.d $L \leq \sqrt{N}$, [23].

5.1 Conclusion

Dans ce chapitre, nous avons présenté et argumenté les résultats des simulations pour différents déploiements de réseau afin d'avoir une vision globale du comportement de notre approche proposée. Nos résultats montrent que la stratégie d'activation/ désactivation basée sur les tabous conjointement avec une allocation d'utilisateurs basée sur le SINR fait en sorte que les BS à fort impact d'interférence choisiraient de s'éteindre et, par conséquent, la décision d'activation/désactivation

est prise en fonction de la rentabilité de ces services. Cependant, notre travail a prouvé qu'il y a beaucoup de variables pouvant susciter des décisions fructueuses ou stationnaires, telles que la densité de l'utilisateur, la longueur du tabou et le poids w_p . En effet, si la taille de la mémoire taboue est exiguë et la densité UE/cellule est grande, la recherche devient répétitive, et donc, l'espace de solutions optimal devient limité. Vice-versa, la taille de la mémoire taboue est large et la densité UEs/cellule est petite, la recherche devient bloquée autour d'un optimal local. Cependant, les gains en puissance et en débit sont optimisés en fonction de la densité UEs/cellule et la longueur de liste taboue. Et les meilleurs résultats sont obtenus quand on priorise le gain QoS avec un poids $w_p \geq 0.5$ pour un gain en puissance qui dépasse 75% et un gain en QoS minimum de 1% en moyenne.

Chapitre 6

CONCLUSION

Dans e3 mémoire, nous avons disséqué le modèle de puissance des différentes stations de base déployées dans un réseau hétérogène pour mieux adapter une stratégie d'économie d'énergie. Premièrement, nous avons déduit que pendant une longue période de désactivation, nous pouvons supposer en toute sécurité que la consommation électrique d'une station de base éteinte est nulle. Nous avons par la suite proposé une stratégie de commutation on/off de la BS basée sur un algorithme de partitionnement de graph, nommée recherche taboue. Cette méthode explore itérativement l'ensemble des solutions d'un problème d'optimisation et par conséquent, nous a permis de dépasser les limites établies par les processus "Hill Climbing" par le biais de la liste taboue, [20]-[21]. La recherche taboue nous permet d'évaluer différentes solutions possibles tout en empêchant le processus d'être bloqué dans les optimaux locaux grâce à la liste taboue; Elle évite de retourner aux solutions déjà rencontrées. Cependant, nos résultats démontrent que si la taille de la mémoire taboue est large, la recherche devient moins flexible. La même remarque est déduite concernant l'impacte de la densité d'UE par cellule. Les gains en puissance et en débit sont optimisés en fonction de la densité UEs/cellule et la longueur de liste taboue. Dans les différents scénarios analysés, les meilleurs résultats sont obtenus quand on priorise le gain en QoS avec un poids $w_p \leq 0.5$.

BIBLIOGRAPHIE

- [1] Nokia Siemens Networks, “ Networks NS,”2011.
- [2] C. Desset, B. Debaillie and F. Louagie, “Towards a flexible and future proof power model for cellular base stations,” in *Tyrrhenian International Workshop on Digital Communications - Green ICT (TIWDC)*, pp. 1–6, Sept 2013.
- [3] M. Feng, S. Mao and T.Jiang, “Base station on-off switching in 5G wireless networks: approaches and challenges,” *IEEE Trans. Wireless Commun*, vol. 24,May 2017.
- [4] G. Xiaohu, H. Cheng, M. Guizani and T. Han, “5G wireless backhaul networks: challenges and research advances” in *IEEE Network*,vol. 28,2014.
- [5] M. Qian, W. Hardjawana, J. Shi and B. Vucetic, “Baseband processing units virtualization for cloud radio access networks,” *IEEE Wireless Commun. letters*, vol. 4,pp.198-192, 2015.
- [6] G. Honggang, K.Wang, J.Hong and V.C.M Leung,“Energy saving in C-RAN based on BBU switching scheme,” *2016 IEEE Internationa Conference on Network Infrastructure and Digital content.*, 2016.
- [7] B. Zhuang, D. Guo and M.L. Honig “Energy-efficient cell activation, user association and spectrum allocation in heterogeneous networks ,” *IEEE J. Sel. Area Commun*, vol. 34,no. 4,pp 823-831,April 2016.
- [8] S. Dahal, “Effect of different distance measures in result of cluster analysis,” *Aalto University School of engineering*,2015.
- [9] L. Tang, W. Wang, Y. Wang and Q. Chen, “An energy-saving algorithm with joint user association, clustering, and on/off strategies in dense heterogeneous networks,” *IEEE Access*, vol. 5, pp. 12988-13000, 2017.
- [10] S. Samarakon,M. Bennis, W. Saad and M. Latva-aho, “Dynamic clustering and on/off strategies for wireless small cell networks,” *IEEE Trans. on Wireless Commun.*, vol. 15, no. 3, pp. 2164-2178, March 2016.
- [11] E. Aqeeli, A. Moubayed and A. Shami, “Power aware optimized RRH to BBU allocation in C-RAN,” *IEEE Trans. on Wireless Commun.*, vol. 17, no. 2, pp. 1311-1322, Feb. 2018.
- [12] M. Khan, R. S. Alhumaima and H. S. Al-Raweshidy, “Reducing energy consumption by dynamic resource allocation in C-RAN,” *2015 European conference on networks and communications*, 2015.

- [13] O. Onireti, A. Mohamed, H. Pervaiz and M. Imran, "Analytical approach to base station sleep mode power consumption and sleep depth", *IEEE 28th Annual International Symposium*, 2017.
- [14] M.E.J.Newman, "Networks: an introduction," *book*, Oxford University Press Inc.,2010.
- [15] M. Deruyck, W. Joseph and L. Martens, "Power consumption model for macrocell and microcell base stations," *Trans. Emerging Tel. Tech*, 2014.
- [16] B. Shen , Z. Lei , X. Huang and Q. Chen, "An interference contribution rate based small cells On/Off switching algorithm for 5G dense heterogeneous networks," *IEEE Access.*, vol. 6, pp. 2169-3536, 2018.
- [17] R. Alhumaima, "Evaluation of power consumption and trade-offs in 5G mobile communications networks," *Brunel University London, Department of Electronic and Computer Engineering College of Engineering*, 2017.
- [18] M. Toulouse, "La recherche avec tabous et ses applications," *Centre de recherche sur les transports*, 1992.
- [19] G. Fred, "A user's guide to Tabu search," *Annals of Operations research*, 1992.
- [20] Per-Olof Fjallstrom, "Algorithms for graph partitioning: a survey," *Linkoping Electronic Articles in Computer and Information Science*, 1998.
- [21] P. Soriano, M. Gendereau, "Fondement et applications des méthodes de recherche avec tabous," *Centre de recherche sur les transports*, 1994.
- [22] U. Von Lucburg, "A tutorial on spectral clustering," *Max Planck Institute for biological cybernetics*, 2007.
- [23] O. Eunsung, S. Kyuho, K. Bhaskar Krishnamachari, "Dynamic Base station switching on/off strategies for green cellular networks," *IEEE Transactions on wireless communications*, 2013.
- [24] S. Bussi, Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, A. Zappone, "A survey of energy efficient techniques for 5G Networks and challenges ahead," *IEEE Journal on selected areas in communications*, 2016.
- [25] J. uo, Q. Chen, "Reducing power consumption by joint sleeping strategy and power control in delay aware C-RAN," *IEEE Access*, 2015.
- [26] C. Lui, B. Natarajan, H. Xia, "Small Cell Base Station Sleep Strategies for Energy Efficiency," *IEEE Transactions*, 2015.
- [27] T. Sigwele, A. S. Alam, P. Pillai, Y. F Hu, "Evaluating Energy efficient cloud radio access networks for 5G," *2015 IEEE Internationa Conference*, 2015.
- [28] H. Fengxia, Z. Shengjie, W. Jinsong, "Survey of strategies for switching off base stations in heterogeneous networks for greener 5G systems," *IEEE Access*, 2016.
- [29] H. Nabuuma, E. Alsusa, W. Pramudito, "A Load-aware base station switch-off technique for enhanced Energy efficiency and relatively identical outage Probability," *IEEE Access*, 2015.
- [30] K. Son et al, "Base Station Operation and User Association Mechanisms for Energy-Delay Trade-offs in Green Cellular Networks," *IEEE JSAC*, 2011.

- [31] M. Feng, S. Mao, and T. Jiang, "BOOST: Base Station ON-OFF Switching Strategy for Energy Efficient Massive MIMO HetNets," *IEEE INFOCOM 2016*, 2016.
- [32] J. Gong et al, "Base Station Sleeping and Resource Allocation in Renewable Energy Powered Cellular Networks," *IEEE Trans. Commun.*, vol. 62, no. 11, Nov 2014.